Identification of determinants for export factors binding to spliced mRNAs

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Agnieszka Marta Gromadzka

aus Posen, Polen

Köln, 2015

Berichterstatter : PD Dr. Niels H. Gehring

Prof. Dr. Karin Schnetz

Tag der mündlichen Prüfung : 13.01.2015

Abstract

The export of messenger RNAs serve as the final posttranscriptional steps of gene expression in the nucleus. The formation of export-competent mRNPs involves the recruitment of export factors, which facilitate transport of mature, processed mRNPs to the cytoplasm.

This study provides a comprehensive analysis of all mammalian export factors. It shows that a core set of export factors including ALYREF, UAP56, and DDX39 readily associate with spliced RNAs in a cap-dependent manner. In contrast, no stable interaction of the export receptor NXF1 or components of the THO-complex was observed. A detailed analysis of ALYREF showed that UAP56 interaction is dispensable, whereas structural integrity of N-and C-terminal regions are essential for RNA binding. Notably, mutating a short conserved motif in an unstructured region of ALYREF abolished the interaction with RNA and proper localization to nuclear speckles.

This results provide important details of the orchestrated recruitment of export factors during the formation of export-competent mRNPs.

Deutsche Zusammenfassung

Der Export von Boten-RNAs ist der finale posttranskriptionelle Schritt der Genexpression im Zellkern. Die Bildung von Export-kompetenten mRNPs beinhaltet die Rekrutierung von Exportfaktoren, die den Transport von reifen, prozessierten mRNPs in das Zytoplasma vermitteln.

Diese Studie liefert eine umfassende Analyse aller Säugetier Exportfaktoren. Sie zeigt, dass ein Kernsatz von Exportfaktoren, einschließlich ALYREF, UAP56 und DDX39, gespleißte RNAs in einer 5′-Cap abhängigen Weise binden. Im Gegensatz dazu, wurde keine stabile Interaktion des Exportrezeptor NXF1 oder Komponenten des THO-Komplexes mit RNAs beobachtet. Eine detaillierte Analyse von ALYREF zeigte, dass die Interaktion mit UAP56 für die Assoziation mit mRNAs nicht notwendig ist, während die strukturelle Integrität der N- und C- terminalen Bereiche für die RNA-Bindung essentiell sind. Bemerkenswert ist, dass die Mutation eines kurzen, konservierten Motivs in einem unstrukturierten Bereich von ALYREF, die Interaktion mit RNA und die korrekte Lokalisierung in nukleären „Speckles" unterbindet.

Diese Ergebnisse liefern wichtige Details zur organisierten Rekrutierung von Exportfaktoren während der Bildung von Export-kompetenten mRNPs.

Table of Contents

Table of Contents

Table of Contents

Introduction

In metazoan cells, RNA-binding proteins interact with pre-mRNA as soon as it is transcribed. Moreover, additional proteins are recruited onto the fully processed mRNA, forming mRNPs – ribonucleoproteins. The exact composition of various mRNPs later influences the stability and translation of the mRNA in the cytoplasm. This section of my thesis describes the sequential recruitment of proteins involved in the multiple stages of mRNPs assembly, capping, splicing, polyadenylation, 3′ end cleavage and deposition of export factors and their role in quality control.

Assembly of an export-competent mRNP in the nucleus

The assembly of mRNPs is already initiated during transcription. Co-transcriptional binding of RNA-binding proteins is mediated through RNA polymerase II (RNAPII) (Moore & Proudfoot, 2009). At the RNAPII exit channel, where newly synthetized mRNA emerges, lies a relatively unstructured domain – the carboxyterminal domain (CTD) of the largest RNAPII subunit (Cramer, 2004; Meinhart, 2005). The CTD consists of heptad repeat structures ($Y_1S_2P_3T_4S_5P_6S_7$) and number of repeats varies from yeast (26 repeats) to mammals (52 repeats) (Bentley, 2005; Tutucci & Stutz, 2011). Each of the serine residues undergoes reversible phosphorylation (Buratowski, 2009; Egloff & Murphy, 2008). Each phosphorylation event marks a specific stage of transcription. Moreover, the genome-wide studies in yeast

revealed that the phosphorylation status of the CTD is gene-specific and not scaled to the gene size (Kim et al, 2010).

Capping and recruitment of the cap binding complex

Phosphorylation of Ser5 in the CTD of RNAPII (a hallmark for an early transcription elongation) recruits capping enzymes, which synthetize a 7-methyl guanylate (m^7G) cap at the 5′ end of pre-mRNA (Egloff & Murphy, 2008; Komarnitsky et al, 2000). The cap is a structural modification of RNA that consist of m^7G linked with an inverted 5′-5 triphosphate bridge to the first nucleoside of a transcript (schematically cap is represented by: m^7GpppN) (Shuman, 2002; Topisirovic et al, 2011). The cap structure prevents newly synthetized pre-mRNA decay by 5′ exonucleases, as well as facilitates its translation in the cytoplasm and influences translation (Muthukrishnan et al, 1975; Schwer et al, 1998; Schwer et al, 2000).

In yeast, three proteins are involved in capping: Ceg1p (guanyltransferase), Abd1p (methyltransferase) and Cet1p (triphosphatase). Phosphorylation of the RNAPII CTD Ser5 residue independently recruits guanyltransferase and methyltransferase (Bentley, 2005; McCracken et al, 1997a). Moreover, the CTD is implicated in enhancing the enzymatic activity of guanyltransferase by enhancing the enzyme affinity for GTP (reducing K_m) (Ho & Shuman, 1999).

In mammalian cells, only two proteins are required for capping: HCE1 (triphospahatase-guanyltransferase) and HCM1 (methyltransferase) (Shuman,

2002). Similar to yeast, guanyltransferase activity of HCE1 is stimulated by phosphorylation of Ser2 residue in the CTD of RNAPII, which in turn activates triphosphatase domain (Ho & Shuman, 1999; Neugebauer, 2002).

The cap structure is bound by the cap binding complex (CBC) (Izaurralde et al, 1995; Izaurralde et al, 1994; Mazza et al, 2001). The CBC consists of two proteins, CBP80 and CBP20, that specifically bind the methylated cap in a co-transcriptional-manner (Topisirovic et al, 2011; Visa et al, 1996). The CBC is implicated in a variety of post-transcriptional processes, such as splicing (Edery & Sonenberg, 1985; Izaurralde et al, 1994; Konarska et al, 1984), export (Abruzzi et al, 2004; Izaurralde et al, 1995; Stutz et al, 2000; Zhou et al, 2000) and translation (Gonatopoulos-Pournatzis & Cowling, 2014). Therefore, capping is a primary example of a co-transcriptional RNA modification that recruits specific RNA-binding proteins (e.g. CBP80, CBP20 and eIF4E), which later influence the life cycle of mRNPs (Bentley, 2005).

Splicing and recruitment of the exon-junction complex

Most genes in higher eukaryotes contain introns and, therefore, their primary transcripts and pre-mRNAs, need to undergo splicing to produce mRNA. Introns need to be removed with single-nucleotide precision to maintain the open reading frame. Pre-mRNA splicing is a two-step transesterifcation reaction that removes introns and ligates exons to produce a functional mRNA. In brief, this reaction is

carried out by the spliceosome, a complex of over hundred proteins, five snRNPs and snoRNPs (Jurica & Moore, 2003; Staley & Guthrie, 1998; Wahl et al, 2009).

Splicing can occur co-transcriptionally, which was discovered by an observation that most transcripts were already spliced while still associated with the chromatin (Alexander et al, 2010; Carrillo Oesterreich et al, 2010). Moreover, it has been shown that the spliceosome induces pausing of the RNAPII at the 3′ exon, which prolongs the time required for assembly of splicing factors on pre-mRNA before release (Alexander et al, 2010; Neugebauer, 2002).

During the splicing process several classes of proteins are recruited to mRNA: serine-arginine rich proteins (SR proteins), proteins of the exon junction complex (EJC) and heterogeneous nuclear RNPs (hnRNPs) (Dreyfuss et al, 2002; Singh et al, 2012). SR proteins generally bind to pre-mRNA within exons in a sequence-specific manner to promote exon definition by recruitment of early spliceosomal components (Shibuya et al, 2004; Zhang & Krainer, 2007). SR proteins contain one or two RNA-recognition motifs in the N-terminus and a C-terminal SR dipeptide that undergoes dynamic phosphorylation (Long & Caceres, 2009). Many SR proteins accompany mRNAs until they reach the cytoplasm (Huang et al, 2003; Huang & Steitz, 2001; Huang & Steitz, 2005; Huang et al, 2004).

After splicing is completed, CWC22 (a spliceosomal component) recruits exon junction complex components through a direct interaction with eIF4A3 (Buchwald

et al, 2013; Steckelberg et al, 2012). The core EJC consists of four proteins – a helicase eIF4A3, a heterodimer MAGOH-Y14 and Barentsz (BTZ). eIF4A3 is a DEAD box helicase, that binds RNA upon hydrolysis of ATP (Bono et al, 2006) and therefore provides a platform for other EJC components (Linder & Jankowsky, 2011). However, eIF4A3 docks on mRNA and binds MAGOH-Y14 independently of ATP-ase activity. Barentsz (BTZ) joins the EJC in the cytoplasm through a direct bipartite interaction with eIF4A3 (Bono et al, 2006; Gehring et al, 2009). Moreover, BTZ can increase eIF4A3 RNA-binding activity (Ballut et al, 2005).

The EJC will remain associated with mRNA during its export to the cytoplasm and provide a platform for a variety of factors implicated in various cellular processes such as localization, translation and degradation (Chazal et al, 2013; Gehring et al, 2009; Le Hir et al, 2001a; Le Hir et al, 2001b; Le Hir et al, 2000b).

3′ end processing and polyadenylation

Pre-mRNA also undergoes processing at the 3′ end, which involves cleavage and polyadenylation. Similar to the capping enzymes, recruitment of polyadenylation and cleavage factors is influenced by modifications in the CTD of RNA polymerase II. In fact, the CTD of RNAPII is required for the association of cleavage and poladenylation factors with the pre-mRNA (McCracken et al, 1997b). Moreover, *in vitro* 3′ end processing is not dependent on the transcription process, but on the presence of the CTD itself (Hirose & Manley, 1998). Recently, the exact role of the

CTD in 3′ end processing has been elucidated in more detail. In human cells the mutant of RNAPII lacking Ser2 residue are deficient in 3′ end processing, while pre-mRNA can still be transcribed (Gu et al, 2013). The phosphorylation status of Ser2 residue reciprocally influences 3′ end processing. This mechanism occurs through RNAPII pausing, recruitment of kinases phosphorylating Ser2, which in turn promotes recruitment of cleavage factor Cstf77 (Davidson et al, 2014).

Mechanisms of mRNA export

RNAs (e.g. mRNAs, tRNA, miRNAs, snRNAs) are transcribed in the nucleus, transported to the cytoplasm where they fulfill their diverse functions. Among various classes of RNAs mRNAs represent a distinct class of RNAs, which requires a specific export system. In contrast to other classes (e.g. tRNAs, miRNAs), mRNAs lack a specific structure that could be a potential target for a nucleocytoplasmatic receptor. mRNAs differ to a high extend in length, sequence and structure (Köhler & Hurt, 2007). Moreover, in contrast to other RNA classes export receptors, mRNA export transport receptor (NXF1) is unrelated to karyopherins (Reed & Hurt, 2002). Furthermore, transport of export-competent is independent of RanGTP – RanGDP gradient (Conti & Izaurralde, 2001) but influenced by various export adaptors (Köhler & Hurt, 2007). Although mammalian and yeast mRNA export share many homologous adaptor proteins (Strässer et al, 2002), the composition ot these export machineries on mRNPs differ substantially (Figure 1). Splicing is a highly

abundant modification of pre-mRNA in the mammalian system. However, genes in yeast are often intronless, resulting in a different mRNP composition in these two systems.

Export complex in lower eukaryotes – Saccharomyces cerevisiae

The yeast mRNA export complex – TREX (transcription-export complex) consists of a THO complex (multiprotein complex containing: Tho2, Hpr1, Mft1, Thp2, Tex1), two adaptor proteins – Sub2 and Yra1 and an export receptor – Mex67 that functions as a heterodimer together with Mtr2 (Strässer et al, 2002; Zenklusen et al, 2002).

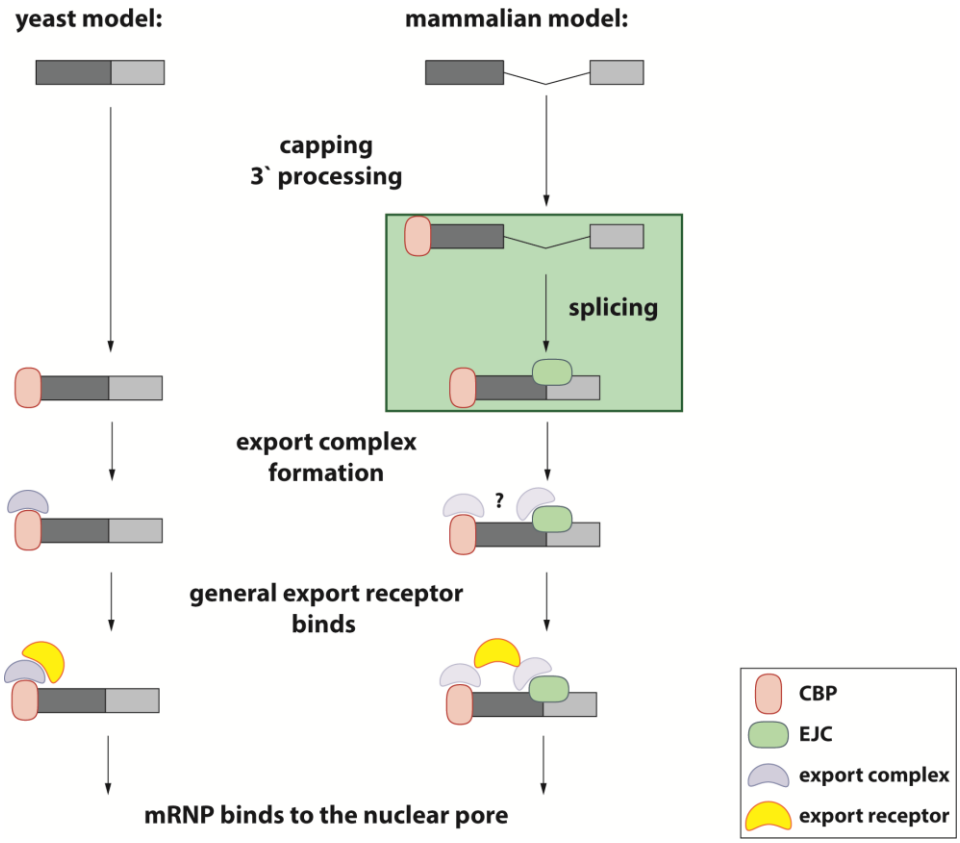


Figure 1. Comparison of export pathways in yeast and mammals.

TREX complex components are linked with various co-transcriptional modifications in yeast. Recently, a direct interaction between Yra1 and the CTD was reported (Meinel et al, 2013). Moreover, Yra1 is also implicated in 3′ end processing through the interaction with Pcf11 (essential component of cleavage and polyadenylation factor IA (CF IA)) (Johnson et al, 2011; Rougemaille et al, 2008). Furthermore, TREX components were found to prevent the accumulation of transcription-mediated obstacles (e.g. R-loops) and function in transcription-coupled DNA repair (Gaillard et al, 2007; Gomez-Gonzalez et al, 2011).

Export complex in higher eukaryotes – mammalian cells

The majority of mammalian genes contain introns, whereas yeast genes are often intronless (Neuvéglise et al, 2011; Shabalina et al, 2010). Therefore, splicing in mammalian cells plays an important role in mRNP remodelling and, thereby, influences gene expression during later steps of the mRNP life cycle. In consequence, the deposition of export components occurs mainly after splicing (Luo et al, 2001; Valencia et al, 2008; Zhou et al, 2000). Exon junction complexes deposited on mRNA during splicing were reported to interact with many of the TREX components (Le Hir et al, 2001b; Le Hir et al, 2000b; Tange et al, 2005). Moreover, EJC deposition greatly enhances export of mRNAs (Zhou et al, 2000).

Core export adaptors

Mammalian mRNA export is mediated by the TREX complex. This complex consist of two adaptor proteins – ALYREF, UAP56 and the THO complex. Adaptor proteins – ALYREF and UAP56 were defined to be involved in mammalian mRNA export based on their yeast homologues – Yra1 and Sub2, respectively (Rondón et al, 2002). Moreover, another export adaptor has been identified through the homology to the yeast adaptor Sub2– DDX39 (Yamazaki et al, 2010). This section describes function of adaptor proteins with existing homologues in yeast system – ALYREF, UAP56 and DDX39.

ALYREF

ALYREF (also referred to as: THOC4, BEF, ALY or REF) is a conserved mRNA export factor. In yeast the ALYREF homologue - Yra1 is an essential factor (Strasser & Hurt, 2000). The depletion of Yra1 in yeast can be complemented with murine ALYREF, however knockdown of ALYREF in mammalian cells has little effect on the export of polyadenylated mRNA (Gatfield & Izaurralde, 2002; Strasser & Hurt, 2000).

Initially, ALYREF was described as EJC component together with SRm160, DEK, RNPS1 and Y14 (Le Hir et al, 2000a; Le Hir et al, 2000c; Le Hir et al, 2000d), which was consistent with the finding that ALYREF binds preferentially to spliced mRNA (Luo et al, 2001).

Mechanistically, ALYREF is recruited to mRNA through a direct interaction with UAP56. This interaction is mediated through short N- and C-terminal regions of ALYREF (Hautbergue et al, 2009; Luo et al, 2001; Zhou et al, 2000). Furthermore, after UAP56 dissociates from the complex, ALYREF binds directly to NXF1 (general export receptor), (Luo et al, 2001; Strasser & Hurt, 2000; Zhou et al, 2000).

Yra1p (ALYREF homologue in yeast) is implicated in influencing 3′end processing (Johnson et al, 2011). Conversely, it has been shown that transcripts with faulty 3′end processing are retained in the nucleus (Saguez et al, 2008). However, in the mammalian system ALYREF is recruited to mRNA co-transcriptionally through the interaction with EJC and a splicing component UAP56 (Cheng et al, 2006; Masuda et al, 2005a). Thus, it still remains elusive what the determinants for ALYREF interaction with mRNPs are and what characteristics of mRNPs are required for the recruitment of ALYREF.

UAP56

UAP56 (56 kDa U2AF-associated protein, also referred to as DDX39B, BAT1) is a DECD-box related RNA helicase containing RNA- and ATP- binding motifs (Shi et al, 2004). UAP56 shares 60% amino acid identity with Sub2 (Figure 2) and can complement loss of Sub2 in yeast (Jensen et al, 2001b; Zhang & Green, 2001). In fruit fly knockdown of HEL (UAP56 homologue in *Drosophila melanogaster*) inhibits growth and leads to nuclear retention of polyadenylated mRNAs (Gatfield et al,

2001). In mammalian cells, UAP56 knockdown was reported to impair mRNA export (Yamazaki et al, 2010), however it was also shown that depletion of UAP56 has no effect on this process (Kapadia et al, 2006).

Mechanistically, UAP56 is involved in the recruitment of ALYREF to mRNA (see section about ALYREF) during mammalian mRNA export. After ALYREF recruitment to mRNA, UAP56 dissociates from mRNA in an ATP-dependent manner (Luo et al, 2001).

Moreover, UAP56 is also an essential splicing factor. It was initially implicated in splicing by its interaction with splicing factor U2AF65 in yeast two-hybrid screen (Fleckner et al, 1997). UAP56 was reported to facilitate the removal of U2AF65 from the polypyrimidine tract downstream of the splice branch point. Thereby, allowing U2 snRNP interaction with the branch point sequence (Kistler & Guthrie, 2001; Libri et al, 2001).

DDX39

DDX39 (alternatively: URH49 – UAP56 related helicase 49 kDa) is mammalian Sub2 homologue. DDX39 shares 90% amino acid sequence identity with UAP56 (Linder & Jankowsky, 2011; Pryor et al, 2004; Shi et al, 2004), (Figure 2). DDX39 binds ALYREF in the same manner as UAP56. Both helicases can complement Sub2 deletion in yeast (Pryor et al, 2004). In mammalian cells knockdown of UAP56 or DDX39 alone has a mild effect on localization of polyadenylated mRNAs. However, combined

knockdown of both helicases leads to nuclear retention of mRNAs (Kapadia et al, 2006). These results suggested that two helicases act in separate export complexes – TREX for UAP56 together with ALYREF and AREX (alternative export complex) for DDX39 and CIP29. Each of the two complexes regulates export of separate target genes (Sugiura et al, 2007; Yamazaki et al, 2010).

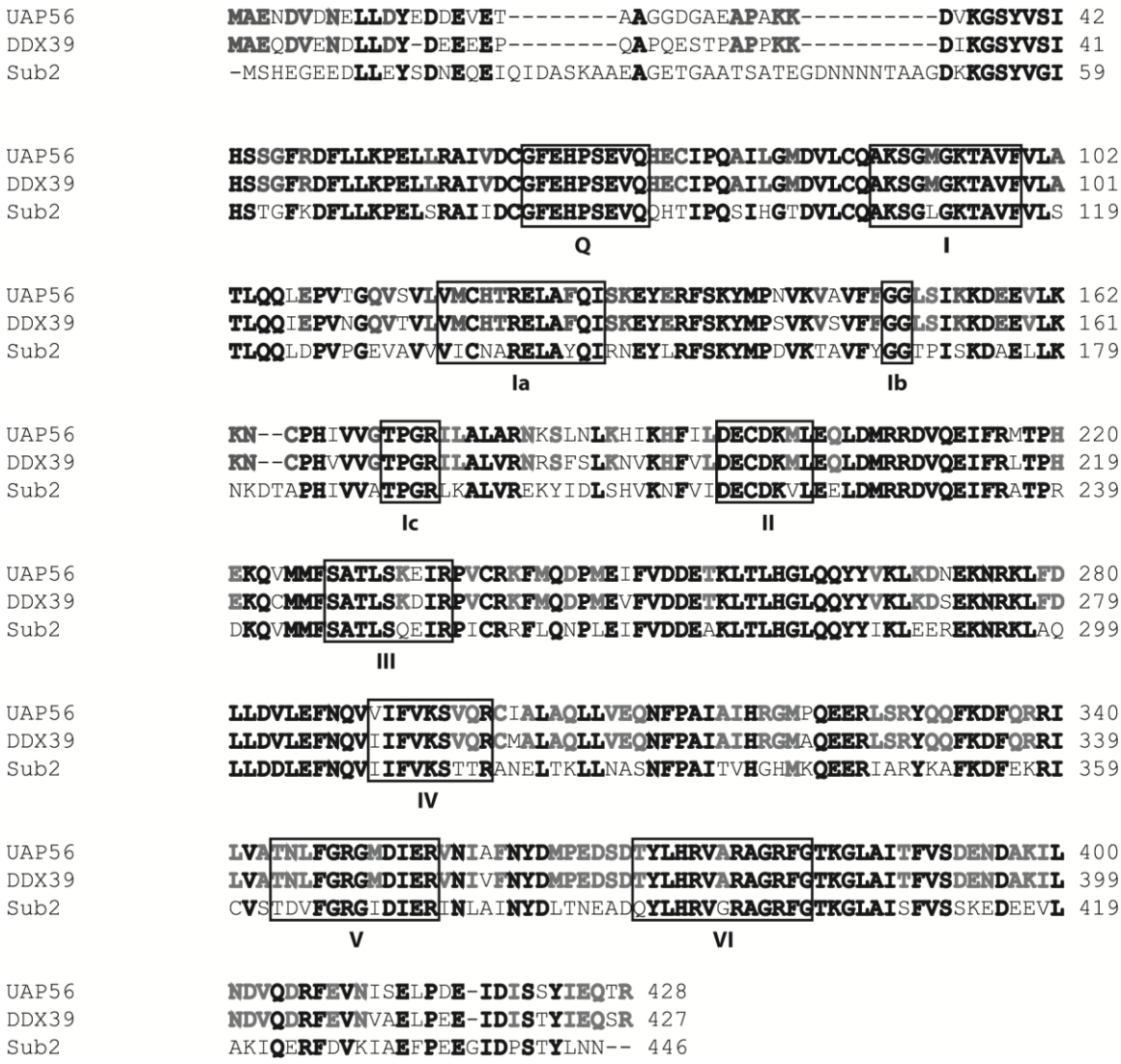


Figure 2. Aligment of yeast Sub2, human UAP56 and human DDX39. Black letters represent amino acids identical for all three proteins, grey letters, identical for UAP56 and DDX39. Black rectangles depict conserved domains.

Additional export adaptors

Apart from ALYREF, UAP56 and DDX39 mRNA export is operated through additional components. The THO complex, another component of the TREX complex is a multiprotein complex that was shown to play a role in UAP56 recruitment. Furthermore, the final step of mRNA export requires the heterodimer NXF1-P15, which binds export-competent mRNA and enables export. Recently, three other components were reported to be important for the process of mRNA export – CIP29, CHTOP and UIF.

THO complex

The mammalian THO complex consists of six components (THOC1, THOC2, THOC3, THOC5, THOC6 and THOC7) (Chi et al, 2013). However, in yeast only four proteins form a THO complex – Hpr1, Tho2 (homologues of the mammalian THOC1 and THOC2, respectively), Mft and Thp2 (Jimeno et al, 2002). Knockdown of each of these four yeast genes results in impaired transcription elongation and genomic stability (Chavez et al, 2000; Piruat & Aguilera, 1998), consitent with the finding that THO complex components are also involved in co-transcriptional 3′ end processing (Rougemaille et al, 2008). Moreover, in yeast the depletion of Tho2 leads to downregulation of the whole THO complex (Masuda et al, 2005b), suggesting that all components are stricktly dependent on the presence of each other. In yeast

strains with defective Mex67 (yeast homologue of NXF1) codepletion of any of the THO components is lethal, suggesting an interplay between THO and Mex67 (Jimeno et al, 2002).

The THO complex components are involved in the deposition of other export adaptors in yeast and in mammalian cells. During yeast export the THO complex recruits Sub2 (UAP56 homologue in yeast) through inteaction with Hpr1 (THOC1 in mammalian cells) (Zenklusen et al, 2002). Furthermore, mammaliann ALYREF interacts directly with THOC2 and THOC5 (Chi et al, 2013). Consistently, THO2 was found to immunoprecipitate spliced mRNA (Cheng et al, 2006). The exact function of the THO complex is not yet understood, although recent studies suggested a role in transcription elongation, 3′ processing and export.

NXF1 and P15

NXF1 (nuclear RNA export factor 1, also referred to as: TAP) is a general export receptor that functions in the cell together with its heterodimerisation partner P15 (also referred to as NXT1 - NTF-2 related protein 1) (Herold et al, 2000; Katahira et al, 1999). NXF1 is a conserved protein, in yeast its homologue Mex67 is required for export of bulk mRNA. Depletion of Mex67 leads to an accumulation of polyadenylated mRNAs in the nucleus. This effect can be complemented by coexpression of human NXF1 and P15. (Braun et al, 2001). Similarly, in fruit fly depletion of NXF1 leads to growth-inhibition and retention of both intronless and

introncontaining, polyadenylated mRNAs (Herold et al, 2001), suggesting a general role of NXF1 in export of all mRNAs.

The studies on NXF1 in mRNA export were initiated by the observation that NXF1 is directly involved in export of a simian type D retroviral constitiutive transport element (CTE) (Kang & Cullen, 1999). In contrast to cellular mRNAs, where NXF1 binds to export-competent mRNAs after several other export adaptors were recruited, NXF1 binds directly to the CTE and this pathway does not require any additional export adaptors (Braun et al, 1999; Grüter et al, 1998).

Although the export of viral CTE element is mediated through a direct RNA-protein interaction, the process of cellular mRNA export is facilitated by the protein-protein interaction of NXF1 and other export adaptor proteins (Liker et al, 2000). Different mechanisms of NXF1 binding to mRNA were proposed. It was suggested that NXF1 has a self-inhibition function. In this model, the RNA-binding activity of RNP-type RNA binding domain (RBD) is suppressed by an intramolecular interaction with the NTF-2 domain of NXF1. It has been suggested that the RBD of NXF1 is released from this suppression and available for RNA-binding, when NXF1 interacts with THOC5 and ALYREF (Viphakone et al, 2012). Interestingly, the presence of RBD was reported to be dispensable for NXF1 binding to mRNA. Furthermore, it was shown that the leucine-rich region (LRR) is required for the interaction with NXF1 (Braun et al, 2001).

The final step of nucleocytoplasmatic translocation is mediated by the components of nuclear pore. NXF1 binds directly through the C-terminal NTF-2 like domain to FG-repeats of the nuclear pore (Bachi et al, 2000; Braun et al, 2002; Fribourg & Conti, 2003). This interaction is stablized by the presense of P15, the heterodimerization partner of NXF1 (Guzik et al, 2001).

CIP29

CIP29 (cytokine-inducible protein 29 kDa) also reffered to as SARNP (SAP domain containng ribonucleoprotein) was initially reported to be involved in growth inhibition and apoptosis induction (Fukuda et al, 2002; Leaw et al, 2004).

Intrestingly the connection between CIP29 and mRNA export was first detected in yeast. Tho1 (yeast homologue of CIP29) was suggested to be involved in an alternative mRNA export pathway, idenpendent from the TREX pathway (Jimeno et al, 2006). This notion was later confirmed in mammalian cells, where it was found that CIP29 together with DDX39 forms an alternative mRNA export complex (see section on DDX39) that functions in parallel to the TREX complex and is involved in export of specific mRNAs during cell cycle (Yamazaki et al, 2010). Furthermore, CIP29 was found to increase the activity of DDX39 in RNA binding and unwinding assay (Sugiura et al, 2007). In contrast, CIP29 was also implicated in the formation of the TREX complex. It was shown that recombinant CIP29 together with UAP56 and ALYREF forms an ATP-dependent trimeric complex (Dufu et al, 2010)

The exact role of CIP29 influencing the formation of TREX complex and its role in the export of specific classes of mRNAs is yet to be understood.

UIF

UIF (UAP56-interaction factor) was initially discovered during a bioinformatic analysis. It was found that ALYREF and UIF share the same UAP56-binding domain (Hautbergue et al, 2009). These predictions were further characterized in the same study. It was reported that, except for UAP56, UIF binds to NXF1 – general export receptor (Hautbergue et al, 2009).

UIF was found to be localized to to nuclear speckles (Hautbergue et al, 2009). Nuclear speckles are dynamic cellular regions where splicing components are stored (Girard et al, 2012). Recently, it was found that also export components (ALYREF, UA56) localize to that cellular region (Dias et al, 2010). The UIF association with mRNA was confirmed using recombinant UIF in mRNP capture and crosslinking assay (Hautbergue et al, 2009).

UIF is suggested to be another export adptor (in addition to CIP29) that influences the composition of mRNPs during assembly of export-cometent complexes. The exact role of UIF remains to be elucidated.

CHTOP

CHTOP (chromatine target of PRMT1) is another newly described export adaptor. Initially this protein was found to be a substrate for the arginine methylase PRMT1 (van Dijk et al, 2010). Simiraly to UIF, CHTOP was first predicted to be involved in mRNA export in bioinformatic analysis. It was found that CHTOP contains the same UAP56-binding motifs found in C-terminus of ALYREF (Chang et al, 2013). The same study confimred that CHTOP binds to UAP56 in mammalian cells. Furthermore, it was discovered that CHTOP binds to NXF1 in a exclusive manner with THOC5 (Chang et al, 2013). Although little is known, CHTOP seems to be another additional export adaptor.

Nuclear retention of mRNAs as a quality control mechanism

The previous sections of this thesis described different steps of mRNA biogenesis that involve deposition of a specific set of proteins to form functional mRNPs. To this end, I have described how proteins required for export, bind to mRNA during each step of mRNP biogenesis (e.g. transcription, capping, 3′ end processing, splicing). Moreover, the main mRNP components, such as EJC components, SR proteins and hnRNPs are also involved in the recruitment of export components. Therefore, all steps of mRNP assembly are important for the formation of so called export-competent mRNPs. These mRNPs are fully processed and can undergo translocation to the cytoplasm where they will be translated into proteins. Hence,

nuclear export and binding of export adaptors are the last nuclear steps of mRNPs assembly and may serve as a quality control mechanism.

During heat shock in yeast all newly transcribed polyadenylated mRNAs are retained in the nucleus, except for heat-shock induced mRNAs. It was shown that heat shock-inducible *HSP104* mRNAs are retained in specific nuclear foci upon knockdown of any of the TREX components, including Mex67 (NXF1 homologue in yeast) (Jensen et al, 2001a). When Mex67 levels are restored (temperature shift for thermosensitive knockout strains), *HSP104* mRNAs from foci are exported and translated in the cytoplasm. This result shows that nuclear foci contain translationally-competent mRNAs retained in the nucleus until exported (nuclear foci do not contain mRNAs awaiting degradation process) .

Retention of heat-shock inducible proteins enables their orchestrated immediate release to the cytoplasm (Kallehauge et al, 2012). Nuclear retention of transcripts (upon restriction of export components) might serve as a control mechanism for tightly regulated expression of a heatshock-inducible expression of proteins.

In yeast the controlled nuclear retention of speficic transcripts upon heat shock is mediated by a protein from a nuclear basket – Mlp1 (Galy et al, 2004). Mlp1 interacts with Mex67 and therefore prevents Mex67 from binding to non-heatshock export-competent mRNPs (Carmody et al, 2010). Moreover, the

interaction between Mlp1 and Mex67 was found to play a role in a retention of intron-containg mRNAs in the cytoplasm (Galy et al, 2004).

Nuclear retention of transcripts during neuronal development enables tightly regulated expression of proteins in a specific localization of neurons. It has been shown that several presynaptic proteins (e.g. tSNARE, vSNARE, Sv2a and Kif5a) are regulated by the nuclear retention of intron-contatining transcripts (Yap & Makeyev, 2013). The mRNAs encoding these neuron-specific proteins are transcribed at detectable levels in both neuronal and non-neuronal cells. However, in non-neuronal cells Ptbp1 represses splicing of 3′-terminal introns of these presynaptic proteins. Retained introns lead to impaired export and retention in the nucleus, where mRNAs are degraded. This process is mediated through Tpr, the mammalian homologue of Mlp1 (Yap et al, 2012). It was shown the knockdown of Tpr increases export of intron-contaning mRNAs through TREX complex (Coyle et al, 2011).

In conclusion, there is a growing evidence that mRNA export might play a substantial role in guiding the expression of heat-shock-related proteins in yeast, as well as in adjusting levels of specific proteins during neuronal development in mammalian cells. However, the mechanisms of mRNA export remain incompletely understood, and therefore it is not yet clear how nuclear retention of specific transcripts could be controlled.

The importance of research

Co-transcriptional recruitment of protein complexes to mRNA orchestrates the formation of mature mRNPs. The order of interactions that lead to the deposition of EJC and CBC components is well understood. However, the sequential interactions that lead to the recruitment of export factors onto mRNPs in higher eukaryotes is not yet fully elucidated.

First question concerns the recruitment of ALYREF and UAP56. In yeast, Sub2 (UAP56 homologue) recruits Yra1 (ALYREF homologue), which leads to the formation of an export-competent mRNPs that can be exported to the cytoplasm. In contrast, the sequence of events leading to ALYREF and UAP56 recruitment to mRNA in mammalian cells is not yet fully elucidated. In particular, the combinatorial role of EJC and CBC in the deposition of TREX components remains to be fully examined.

Second question concerns the recruitment of NXF1. NXF1 is thought to bind mRNPs and direct them to the nuclear pore. However, the direct interaction of NXF1 had only been studied with viral RNAs, but it is not yet clear if this model applies to cellular mRNAs. Moreover, the role of export adaptors in NXF1 recruitment is not fully understood.

Finally, the influence of the THO complex and the recently discovered export adaptors CIP29, UIF and CHTOP on the assembly of export-competent mRNPs remained to be examined.

Taken together, there is a need for a comprehensive analysis of mRNA recruitment of mammalian mRNA export components. This needs to be examined in a controlled system, in order to propose a model of mammalian mRNA export dynamics.

Research goals

1. To elucidate how the cap structure and splicing-dependent EJC deposition contribute to the binding of export components.
2. To understand the role of THO complex components and additional export components in mRNP assembly
3. To decipher the mechanism of NXF1 recruitment to mRNPs
4. To describe the mutual interplay between export adaptors – ALYREF and UAP56 in the mRNP assembly

Results

The aim of my thesis was to understand how the mRNP composition influences export of mRNA to the cytoplasm in mammalian cells. One of the distinct post-transcriptional modifications of mRNPs in mammalian cells is splicing. Therefore, first four parts address the binding of eleven export factors to various RNA substrates, and the role of splicing in mRNP assembly. The fifth part concerns the visualization of mRNPs in living cells in order to understand the dynamics of export process in the cell.

Recruitment of core export components to mRNPs during splicing

In the first part of my work, I systematically analysed the interactions of proteins involved in mammalian mRNA export (ALYREF, UAP56, DDX39, NXF1, the THO complex components, CHTOP and CIP29) with mRNA substrates. I utilized a well-established *in vitro* splicing system (Gehring et al, 2009; Steckelberg & Gehring, 2014). This system recapitulates binding of various mRNP components to intron-containing (spliced) or intronless (unspliced) mRNA reporter substrates. In this assay, HeLa nuclear extracts were supplemented with splicing-competent HEK293 whole-cell extracts expressing FLAG-tagged export components and *in vitro* transcribed RNA substrate. After splicing FLAG-tagged RNA-protein complexes were immunoprecipitated, RNA was resolved on a denaturing gel and analysed.

Mammalian export components bind specifically to spliced mRNAs

First, I recapitulate binding of export components (DDX39, UAP56, ALYREF and NXF1) and two known mRNP components – CBP80 and SELOR to intron-containing MINX RNA (Figure 3). CBP80 is a cap-binding protein which immunoprecipitates with equal efficiency spliced and unspliced cap-containing RNAs (Izaurralde et al, 1994). SELOR is an eIF4A3-interacting domain of an EJC component – Barentsz that immunoprecipitates spliced mRNA (Degot et al, 2004; Gehring et al, 2009).

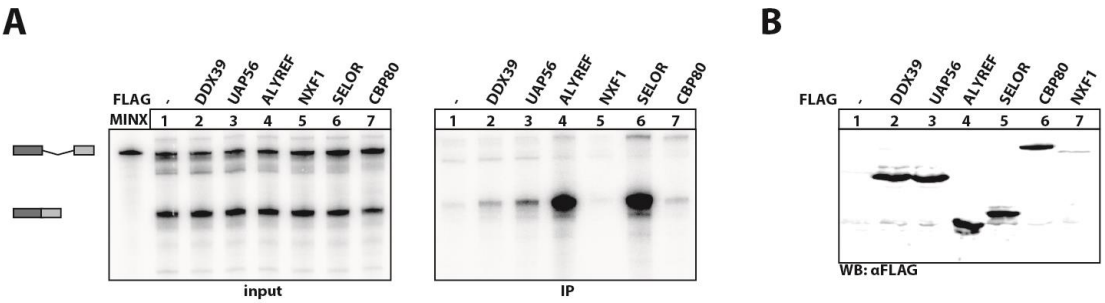


Figure 3. Core export adaptors bind to spliced mRNA *in vitro* (A) *In vitro* splicing reactions of MINX substrate were supplemented with extracts expressing the indicated FLAG-tagged export factors. SELOR (eIF4A3-binding domain of BTZ) and CBP80 were used as controls, unfused FLAG-tag served as a negative control. FLAG-containing mRNPs were immunoprecipitated with anti-FLAG beads and co-precipitated RNA resolved on a denaturing PAGE. 10% of each splicing reaction was used as input. Schematic representations of splicing products are depicted on the left side of the autoradiograph. **(B)** Expression of FLAG-tagged proteins in HEK 293 extracts used in Figures 3-6 was determined by immunoblot analysis using a FLAG antibody.

In this assay ALYREF binds efficiently to spliced mRNA and precipitates comparable amounts of spliced mRNA as SELOR (Figure 3A, lane 4 and 6). The export adaptors UAP56 and DDX39 bind spliced RNAs to a lesser extent than ALYREF (Figure 3A,

lane 2 and 3). Only a weak binding of all these export factors to the unspliced pre-mRNA was observed. This observation is coherent with previous reports (Luo et al, 2001; Masuda et al, 2005b) where it was shown that the process of splicing enhanced binding of export components to mRNA. No interaction of NXF1 with reporter MINX RNA was detected. This observation will be addressed further in a separate section.

Mammalian export components binding to mRNA is enhanced by splicing

To exclude a possibility that the binding of export components to a reporter MINX RNA is mediated by a specific sequence or structure, MINX mRNA without intron sequence (MINX Δi) was used. In this assay the analysed proteins bind to a reporter RNA in a splicing-independent manner.

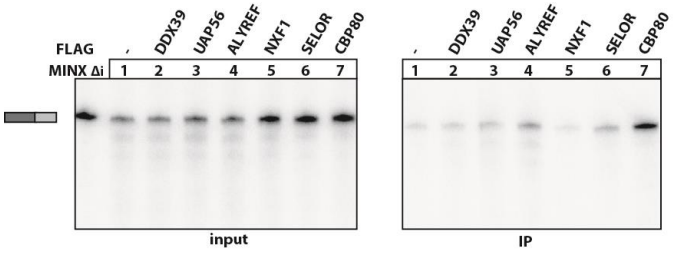


Figure 4. Core export adaptors binding to mRNA is enhanced by splicing. *In vitro* splicing reactions were performed as in Figure 3A using intronless MINX substrate (MINX Δi).

All tested export components (ALYREF, UAP56 and DDX39) precipitated small amounts of MINX Δi, hence export components bind to RNA in a splicing-dependent manner (Figure 4, lanes 2-4). In contrast, CBP80 efficiently assembled into

complexes with cap-containing RNA independently of splicing machinery (Figure 4, lane 7).

Hence, splicing and deposition of post-splicing mRNP components (such as EJC) greatly enhances association of export components to mRNA.

ALYREF and UAP56 binding to mRNA is cap-dependent

Binding of export factors to the cap or the cap region of mRNAs has been previously demonstrated, but the detailed mechanism of this interaction was not fully explored (Cheng et al, 2006). In order to understand the cap-dependent binding of export components, I utilized a modified splicing assay. In this assay capped MINX pre-mRNA was spliced in the excess of cap analogue (m^7GpppG) (Figure 5).

In contrast to previous reports (Edery & Sonenberg, 1985; Konarska et al, 1984; Nojima et al, 2007), the addition of cap analogue to the *in vitro* splicing reaction did not decrease splicing efficiency (Figure 5). In this assay CBP80 and SELOR were used as positive and negative controls, respectively. CBP80 binding to MINX mRNA was decreased in the presence of cap analogue (compare lane 6 and 7 in Figure 5A and 5B), whereas SELOR was not affected (compare lane 4 and 5 in Figure 5A and 5B). Strikingly, binding to spliced RNA of both ALYREF and UAP56 was strongly impaired (compare lane 2 and 3 in Figure 5A and 5B).

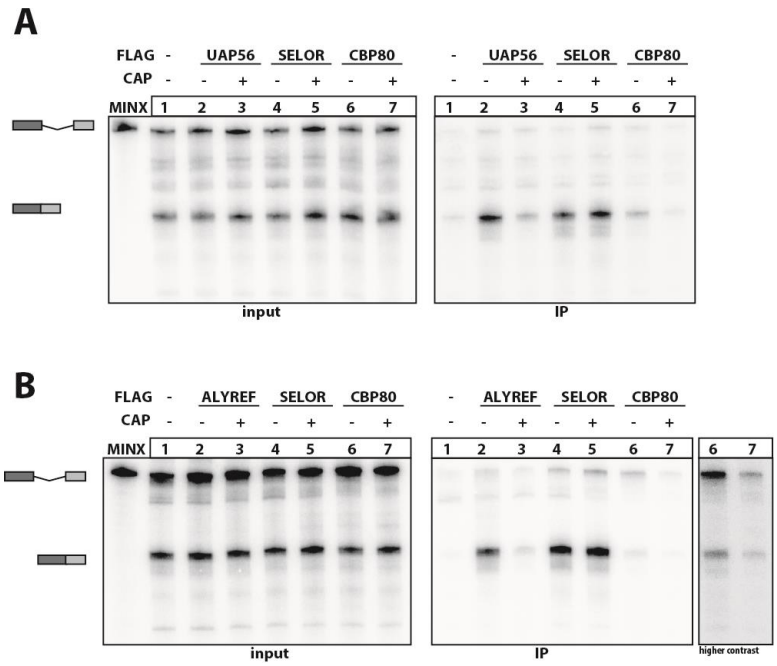


Figure 5. ALYREF and UAP56 binding to spliced mRNA is cap-dependent. *In vitro* splicing reactions were performed as in Figure 3 and supplemented with 10 µM of cap analogue (m^7GpppG) as indicated. mRNPs were immunoprecipitated and RNAs resolved on denaturing PAGE. FLAG-tagged SELOR and CBP80 served as negative and positive controls, respectively. A high-contrast picture of lanes 6 and 7 is shown as separate panel.

The results obtained are in line with the previously reported recruitment of the mRNA export machinery to the 5' end of mRNA (Cheng et al, 2006). However, here I show that binding of ALYREF and UAP56 to spliced mRNA requires cap-binding proteins.

Export factors binding to RNA requires the presence of both the cap and the EJC

I found that export components bind to spliced mRNA and that this binding requires presence of the cap structure. To determine the position where the core export components bind on the RNA, I used oligonucleotide-directed RNase H activity. The

oligonucleotide sequence that directs RNAse H to a specific position was designed in order to cleave the RNA after splicing in a region upstream of the EJC-binding site, in a first exon, which enables to separately examine the requirement of the cap or EJC in the recruitment of export factors. Each fragment of cleaved RNA represents a specific characteristics of mRNPs. The uncut spliced mRNA contains both the cap and EJC, the 5' fragment of the spliced mRNA contains only the cap, but no EJC, and the 3' fragment of spliced mRNA contains EJC, but not the cap (Figure 6). A similar assay was previously used to analyse the binding of the TREX complex to the spliced RNAs (Cheng et al, 2006).

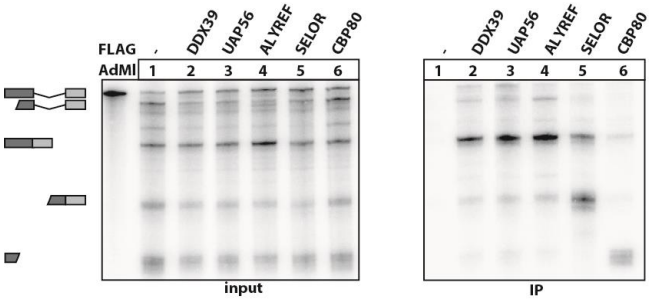


Figure 6. Core export component bind to intact, spliced and capped mRNA *in vitro*. *In vitro* splicing reactions were performed as in Figure 3 using AdMI ex2 substrate followed by incomplete RNase H digestion of splicing products and immunoprecipitation of mRNPs. Schematic representations of splicing and digestions products are depicted on the left side of the autoradiograph.

Although DDX39, UAP56, and ALYREF efficiently bound to full length spliced RNA, neither of the two fragments produced after RNAse H digestion was efficiently precipitated (Figure 6, lanes 2, 3 and 4). In contrast, SELOR precipitated two fragments with the EJC binding site, 3' fragment and the full length spliced mRNA

(Figure 6, lane 5). The cap-binding protein CBP80, which precipitated the cap-containing 5′ fragment (Figure 6, lane 6), showed their expected binding preferences. Hence, export factors may require the presence of both, the cap and the EJC, to establish a stable interaction with mRNA.

Additional export components moderate mRNP composition during splicing

In addition to the core export components (e.g. ALYREF, UAP56, DDX39, NXF1), I also examined additional factors that were described to be involved in the formation of export-competent complexes. In the following part, I tested THO complex components, CHTOP and CIP29 for their binding to the mRNA during splicing.

THO complex components binding to mRNA during splicing

THO complex consist of six proteins: THOC1, THOC2, THOC3, THOC5, THOC6 and THOC7. Binding of THOC2 and THOC5 to the spliced mRNA *in vitro* has been reported previously (Cheng et al, 2006; Masuda et al, 2005b). Therefore, it was important to assess contribution of THO complex in mRNP formation during splicing.

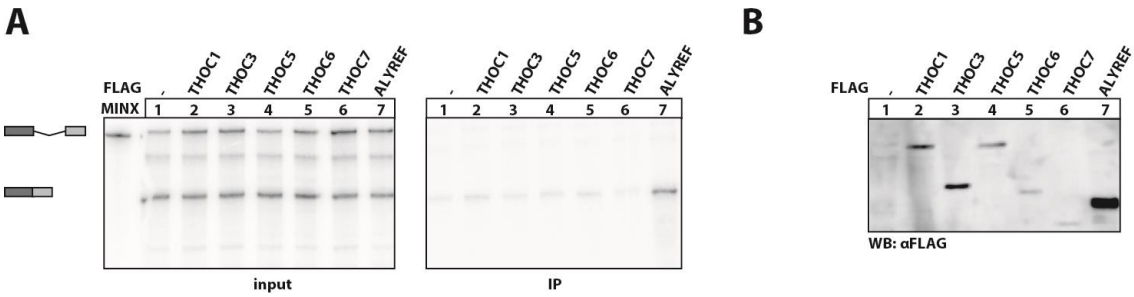


Figure 7. THO complex binding to mRNA during splicing. (A) *In vitro* splicing reactions of MINX were supplemented with extracts expressing FLAG-tagged THO complex components and performed as described in Figure 3. **(B)** Expression of FLAG-tagged proteins in HEK 293 extracts used in panel A was determined by immunoblot analysis using a FLAG antibody.

No RNA binding to any of the tested components of the THO complex (THOC1, THOC3, THOC5, THOC6 and THOC7) was observed (Figure 7). ALYREF was used as a positive control of a protein which binds specifically to the spliced mRNA. In this study, THOC2 was not analysed due to its low expression levels. Although binding of the THO complex proteins to RNA has been previously reported (Cheng et al, 2006), however, it was not possible to reproduce their stable interaction with spliced RNA in this thesis. Therefore, the exact contribution of THO complex components to mRNPs remains to be elucidated.

Auxiliary factors

Recently additional export factors were found to be required for mRNA export – CIP29, CHTOP and UIF. In this study binding of CIP29 and CHTOP to MINX RNA was assessed. Consistently, it was not possible to generate lysates overexpressing UIF

due to the undetectable levels of this protein. Therefore, the impact of this protein on mRNP formation is not investigated in this study.

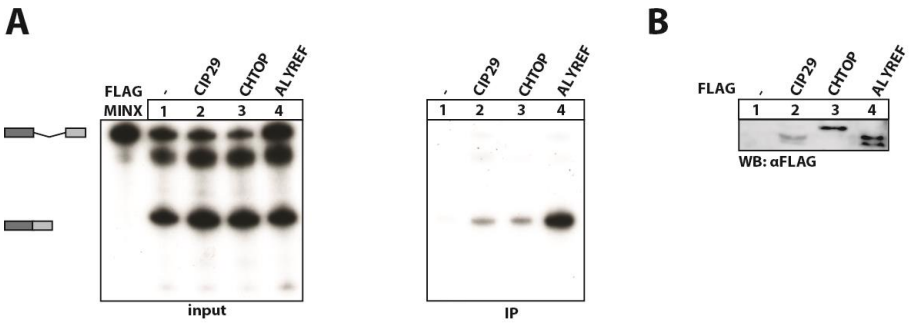


Figure 8. Binding of auxiliary export factors to mRNA during splicing (A) *In vitro* splicing reactions of MINX were supplemented with extracts expressing FLAG-tagged CIP29 and CHTOP and performed as described in Figure 3. **(B)** Expression of FLAG-tagged proteins in HEK 293 extracts used in panel A was determined by immunoblot analysis using a FLAG antibody.

Both CIP29 and CHTOP bind spliced mRNA (Figure 8, lane 2 and 3), although to a lesser extent than ALYREF (Figure 8, lane 4) that serves here as a positive control. Hence, due to the minor amounts of precipitated mRNA by CIP29 and CHTOP further analysis of the cap or the EJC involvement in the binding process was not feasible. Both CIP29 and CHTOP bind to the mRNA during splicing, however these results suggest that auxiliary factors and core export components (e. g. UAP56, DDX39, ALYREF) interact transiently, what can be observed by a small amounts of precipitated mRNA in a splicing assay. The exact mechanism of this process remains to be elucidated in more detail.

NXF1 binds differentially to various RNAs *in vitro*

NXF1 is a general export receptor that binds all processed mRNPs and the nuclear pore, hence enabling mRNA translocation to the cytoplasm. Therefore, it was surprising that no interaction between MINX mRNA and FLAG-tagged NXF1 was detected (Figure 3, lane 5).

NXF1, unlike other export components, binds to CTE

NXF1 consist of a RNP-type RNA binding domain (RBD). However, its interaction with RNA has been reported to be suppressed by an intramolecular interaction with the NTF-2 domain of NXF1. It has been suggested that the RBD is released from this suppression and is available for RNA-binding, when NXF1 interacts with THOC5 and ALYREF (Viphakone et al, 2012).Although, for export of viral RNAs a different mechanism was proposed. During export of viral RNAs, NXF1 was found to directly bind CTE (constitutive transport element), an RNA originating from a simian type D retrovirus, this direct binding does not require any additional factors for an efficient transport of viral RNA to the cytoplasm (Bachi et al, 2000; Grüter et al, 1998).

Initially to recapitulate NXF1 binding to RNA, I utilized CTE RNA (from Mason-Pfizer monkey virus) for *in vitro* binding assays with export components (Figure 9). I used the same experimental strategy as for *in vitro* splicing assay. Briefly, HeLa nuclear extracts were supplemented with whole-cell extracts expressing FLAG-tagged

export components. Capped, intronless CTE RNA serves as a reporter RNA for a binding assay used in this study.

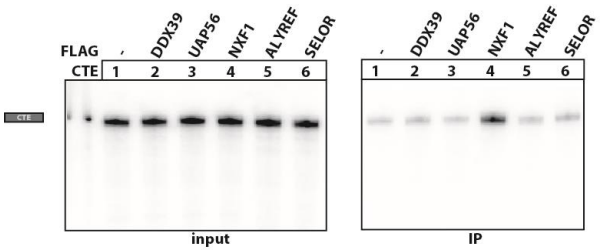


Figure 9. Export components binding to viral CTE. *In vitro* binding reaction of CTE RNA to FLAG-tagged TREX components. RNA-protein complexes were immunoprecipitated and co-precipitated RNA resolved on a denaturing gel (as in Figure 3). 10% of each reaction served as input.

In Figure 9, I observed efficient precipitation of CTE RNA by FLAG-NXF1, demonstrating that the NXF1 can bind to RNA and is in principle functional. In contrast, none of the other export factors (DDX39, UAP56 or ALYREF) bound to the CTE RNA. Hence, in the further *in vitro* assays I used CTE RNA to study the molecular details of RNA-binding of NXF1 and the function of its self-inhibitory domain.

NXF1 binding to CTE is promoted by the removal of RBD inhibitory mechanism

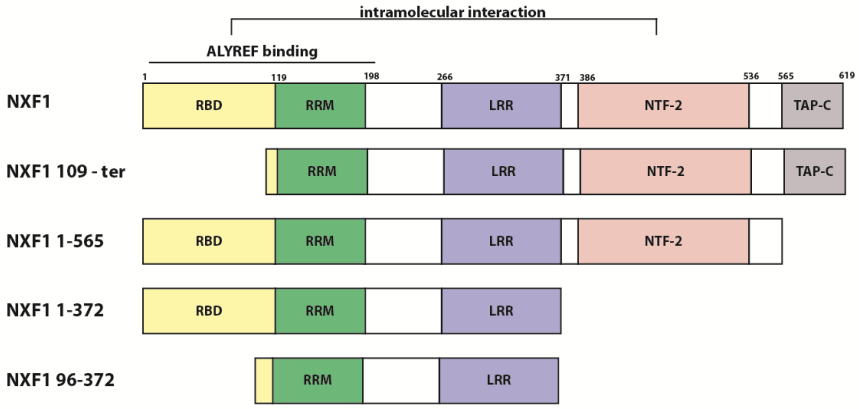
In order to understand the discriminants for NXF1 binding to CTE, series of truncation mutants were tested (Figure 10A). Mutants lacking NTF-2 domain (NXF1 1-372 and 1-565) are no longer able to sustain intramolecular, RBD-inhibitory mechanism. C-terminal TAP-C domain is required for interaction with FG-repeats of nucleoporins, shuttling activity and localization to the nuclear rim (Bachi et al, 2000; Bear et al, 1999; Braun et al, 2001; Herold et al, 2000). An additional N-

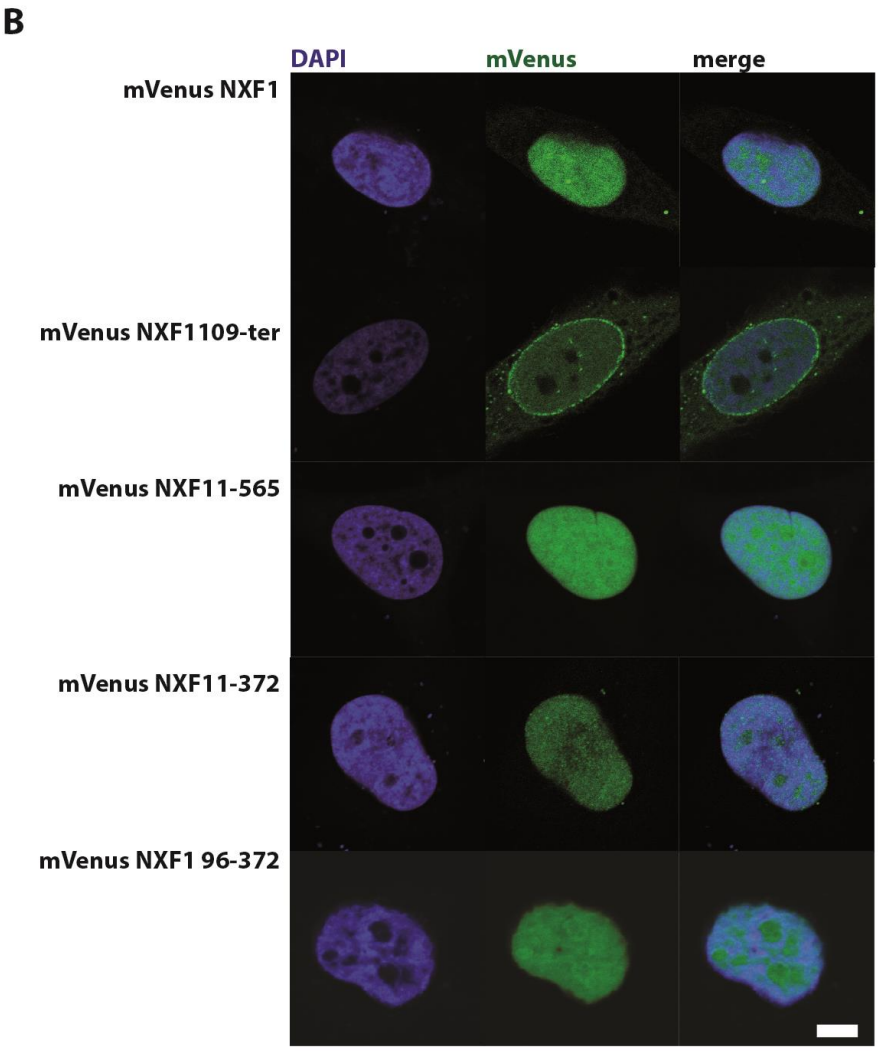
terminal (1-105 aa) arginine-rich motif was found to be important for RNA binding (Hautbergue et al, 2008).

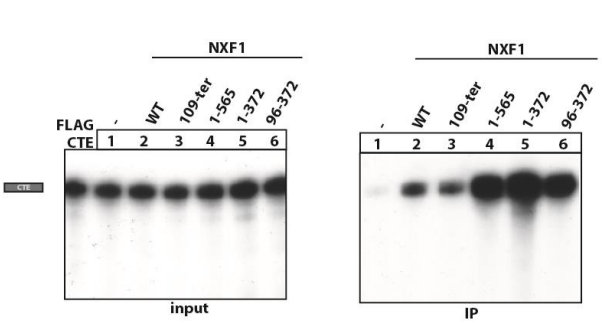
Fusion constructs of all NXF1 mutants with mVenus were constructed in order to test their functionality (Figure 10B). All tested NXF1 truncations localized, as the full length protein to the nucleus, except for 109-ter deletion. In the case of the NXF1 109-ter NLS was deleted, however fusion protein localize to the nuclear rim. The TAP-C domain is required for this feature (Bachi et al, 2000).

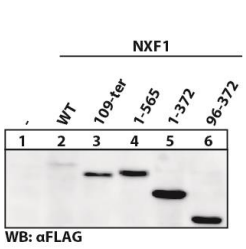
Next, all NXF1 mutants were used in the CTE binding assay. All NXF1 mutants precipitated the CTE RNA. However, mutants NXF1 1-372 and NXF1 1-565 (lacking self-inhibitory mechanism) showed an increased binding to the CTE (Figure 10C, lanes 4 and 5), indicating that permanently exposed RBD increases association of NXF1 with the CTE.

A




B
DAPI
mVenus
merge
mVenus NXF1
mVenus NXF1109-ter
mVenus NXF11-565
mVenus NXF11-372
mVenus NXF1 96-372


C
NXF1
FLAG
CTE
-
WT
109-ter
1-565
1-372
96-372
1 2 3 4 5 6
input
IP


D
NXF1
-
WT
109-ter
1-565
1-372
96-372
1 2 3 4 5 6
WB: αFLAG

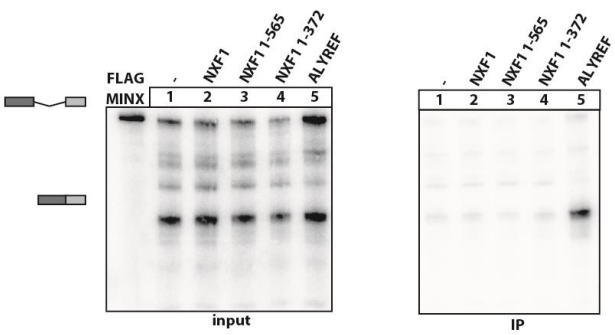

E
FLAG
MINX
-
NXF1
NXF1 1-565
NXF1 1-372
ALYREF
1 2 3 4 5
input
IP

Figure 10. NXF1 binding to various RNA targets *in vitro* (A) Schematic representation of NXF1 deletion mutants used in this study. **(B)** Localization of NXF1 mutants in HeLa cells. DAPI was used to stain nuclear DNA. Scale bar, 5 µm. **(C)** *In vitro* binding reaction of CTE RNA to FLAG-tagged NXF1 mutants. **(D)** Expression of FLAG-tagged proteins in HEK 293 extract used in panel C. **(E)** *In vitro* splicing reactions of MINX substrate supplemented with extracts expressing FLAG-tagged NXF1 truncation mutants (as in Figure 1). FLAG-tagged ALYREF served as a positive control. Schematic representations of splicing and digestions products are depicted on the left side of the autoradiograph.

Furthermore, I re-examined C-terminal deletion mutants of NXF1, lacking the NTF-2 domain – FLAG-tagged NXF1 1-372 and NXF1 1-565 in splicing assay (Figure 10E) where MINX RNA was used as a substrate. It was not possible to stabilize the interaction of the general export receptor NXF1 mutants lacking RRM-inhibitory mechanism and spliced MINX RNA. Therefore, it is likely that mRNPs only transiently recruit NXF1 shortly before or during the export process.

NXF1 binding to CTE is promoted by P15

In mammalian cells, NXF1 binds P15 and forms a functional heterodimer. Both proteins contain a NTF-2-like domain, that mediates the interaction between the heterodimerization partners (Braun et al, 2001). Moreover, in fruit fly both P15 and NXF1 are essential for efficient export of polyadenylated mRNAs (Herold et al, 2001). Here I sought to test binding of each heterodimer components in CTE binding (Figure 11).

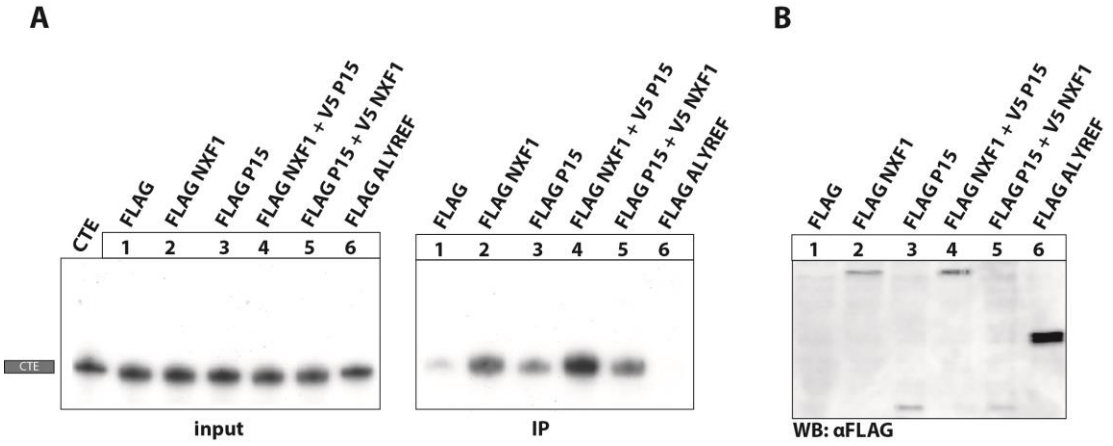


Figure 11. NXF1 and P15 binding to CTE RNA. (A) *In vitro* binding reaction of CTE RNA to FLAG-tagged NXF1 and P15 alone or with V5-tagged heterodimerisation partner. RNA-protein complexes were immunoprecipitated and co-precipitated RNA resolved on a denaturing gel (as in Figure 3). 10% of each reaction served as input. **(B)** Expression of FLAG-tagged and V5-tagged proteins in HEK 293 extracts used in (A) was determined by immunoblot analysis using a FLAG antibody and V5 antibody (respectively).

The heterodimer NXF1:P15 binds to RNA (Katahira et al, 1999), therefore in lysates RNA binding of one of the heterodimer components is limited by the expression of the other. To overcome this limitation, both heterodimer components were overexpressed with FLAG-tag and V5-tag, respectively (Figure 11). When FLAG-tagged NXF1 and V5-tagged P15 were overexpressed, FLAG-NXF1 immunoprecipitated more RNA than when expressed alone (Figure 11A, compare lane 2 and 4). Notably, the expression of FLAG-tagged NXF1 was higher when V5-P15 was overexpressed (Figure 11B, compare lane 2 and 4). The same observation was true for FLAG-tagged P15 (Figure11A, compare lane 3 and 5), although in general levels of precipitated CTE were lower compared to NXF1.

NXF1 binding to CTE is promoted by the presence of P15, suggesting that both proteins are involved in cooperative binding to the viral element. In consistency, previous reports show that overexpression of both heterodimer components promoted export of transcripts that were inefficiently transported under normal conditions (Braun et al, 2001; Herold et al, 2001). The exact role of P15 in binding to the CTE element is not yet understood and requires further analysis.

NXF1 in vivo binding (CLIP experiments)

NXF1 is a general export receptor that binds all mRNAs allowing translocation to the cytoplasm. To this end here only a direct interaction of NXF1 with viral element CTE was recapitulated (Figure 10C), whereas no binding between NXF1 and MINX mRNA was detected (Figure 3, Figure 10E), likely due to a transient interaction of NXF1 with export competent mRNPs.

In order to recapitulate NXF1 binding to mRNA in living cells, I used the CLIP (UV crosslink and immunoprecipitation) assay (Huppertz et al, 2014; Konig et al, 2011). Briefly, in this assay HEK-293 cells expressing FLAG-tagged NXF1 proteins upon induction (or truncation mutants) were induced with doxycycline, crosslinked with 400 mJ UV light and harvested. mRNAs covalently bound to a protein are subjected to limited RNAse I digestion. The limited digestion allows to stabilize mRNA fragments suitable for further labelling and later detection. Afterwards, FLAG-tagged protein:mRNA complexes are co-immunoprecipitated. Polynucleotide

kinase reaction with αP^{32}-ATP radioactively-labels mRNA fragments crosslinked to proteins. Proteins are resolved on a gel and radioactive signals are analysed. The next step of CLIP assay requires deep-sequencing of recovered RNA fragments and bioinfomatic analysis.

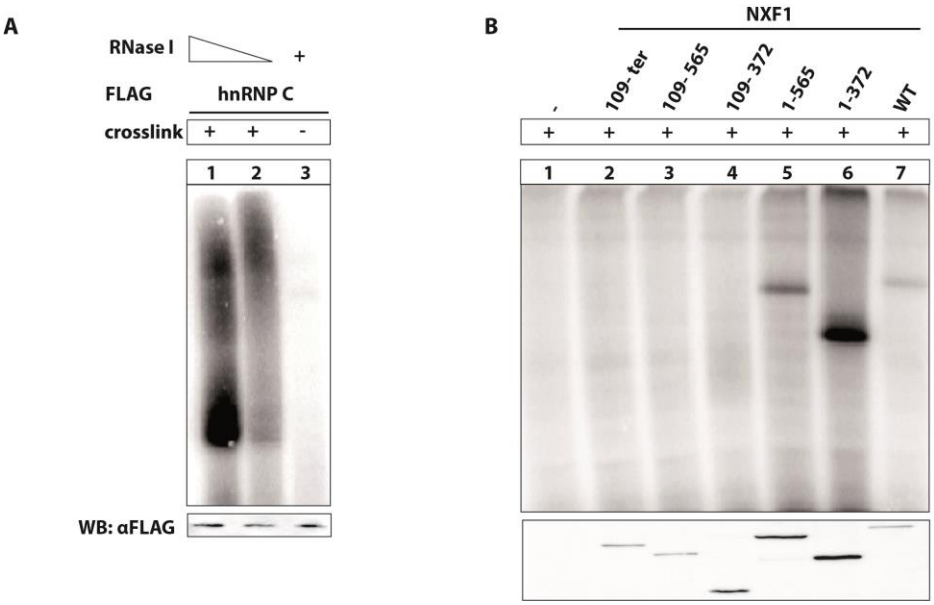


Figure12. NXF1 *in vivo* binding to mRNA. (A) Lysates from Flp-In T-REx HEK-293 cell lines expressing FLAG hnRNP C were crosslinked (+) or not (-) and digested with different RNAse I concentrations, crosslinked mRNA was labelled with PNK reaction. **(B)** Lysates from FlpIn T-REx HEK-293 cell lines expressing different FLAG NXF1 variants or unfused FLAG peptide (as a control) were crosslinked (+).

Initially, crosslinking was re-examined for FLAG-tagged hnRNP C (Figure 12A). hnRNP C was previously described to efficiently crosslink to RNA in a CLIP assay (König et al, 2010). The immunoprecipitated complex of RNA and protein was present in UV-crosslinked samples and absent when UV-crosslinking was omitted (Figure 12A, compare lane 3 with 1 and 2). Two different RNase I concentrations

were tested. In high RNase I concentration short RNA fragments were generated and relatively distinct band on a gel corresponding to the size of a protein was present (Figure 12A, lane 1). In contrast, after limited digestion, in low concentration of RNase I, long fragments were predominant (Figure 12A, lane 2), leading to the up-shift of radioactive signal.

Next, I applied the CLIP assay on FLAG-tagged NXF1 wild type and truncation mutants. The FLAG-tagged NXF1 N-terminal truncation mutants did not form mRNA:protein complexes upon crosslining, no specific signals were observed on a gel (Figure 12B, lane 2, 3 and 4). In contrast, wild type FLAG-tagged NXF1 and C-terminal deletions formed mRNA:protein complexes (Figure 12B, lane 5, 6 and 7), what appeared on a gel as a sharp band corresponding to the size of a protein. The absence of crosslink signals for N-terminal deletions can be explained by the lack of nuclear localization of mutants lacking NLS (Figure 13). Therefore, only mutants that are physically localized to nucleus could be crosslinked.

The CLIP assay confirmed an interaction of NXF1 and mRNA. However, the low levels of crosslinked FLAG NXF1:mRNA complexes in comparison to the crosslinked hnRNP C indicated, that next steps of the CLIP assay (including RNA recovery and sequencing) are not feasible.

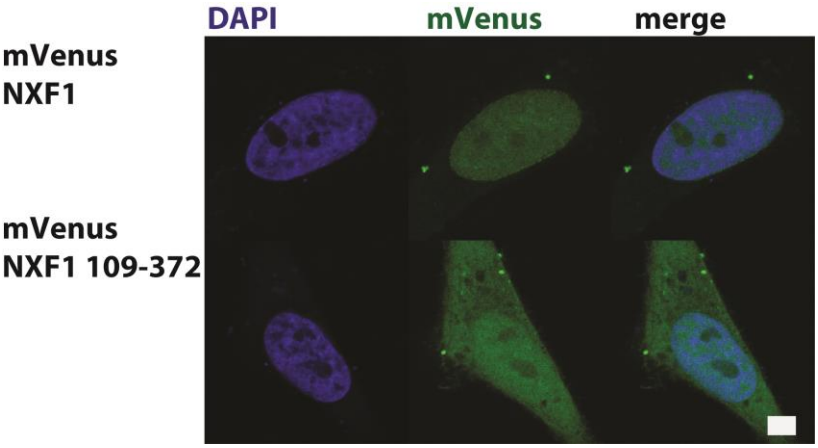


Figure 13. NXF1 lacking N-terminal domain is localized to the cytoplasm. Localisation of mVenus NXF1 and mVenus NXF1 109-ter in HeLa cells. DAPI was used to stain nuclear DNA. Scale bar, 2 µm.

ALYREF binding to mRNA is independent of UAP56 interaction

ALYREF and UAP56, core components of TREX complex efficiently precipitate intact spliced mRNA in a cap-dependent manner. In this section, I elucidate the specific interplay between these proteins and their requirements for recruitment to RNA.

ALYREF truncation mutants, deletion of potential RNA-binding domains

ALYREF contains an RRM (RNA-recognition motif, 106 – 187 aa), two previously described UAP56-interaction domains (N-terminal 1-15 aa, C-terminal 237-257 aa), an NXF1-binding motif (15-37 aa) and two arginine-rich regions (17-87 aa; 198-231 aa) (Luo, Zhou et al. 2001, Golovanov, Hautbergue et al. 2006, Hautbergue, Hung et al. 2009). Notably, the N- and C-terminal domains, but not the RRM, were

sufficient for murine ALYREF binding to RNA *in vitro* (Stutz, Bachi et al. 2000, Rodrigues, Rode et al. 2001).

In order to understand the recruitment of ALYREF to mRNA, truncation mutants of potential RNA-binding domains were generated (N-terminal RGG-rich region, RRM, C-terminal RGG-rich region) (Figure 14A). Truncation mutant ALYREF 1-106 contains only the N-terminal RGG-rich region and the RRM. ALYREF RRM contains just the RNA-recognition motif. ALYREF 106-ter contains C-terminal RGG-rich region and the RRM (Figure 14A).

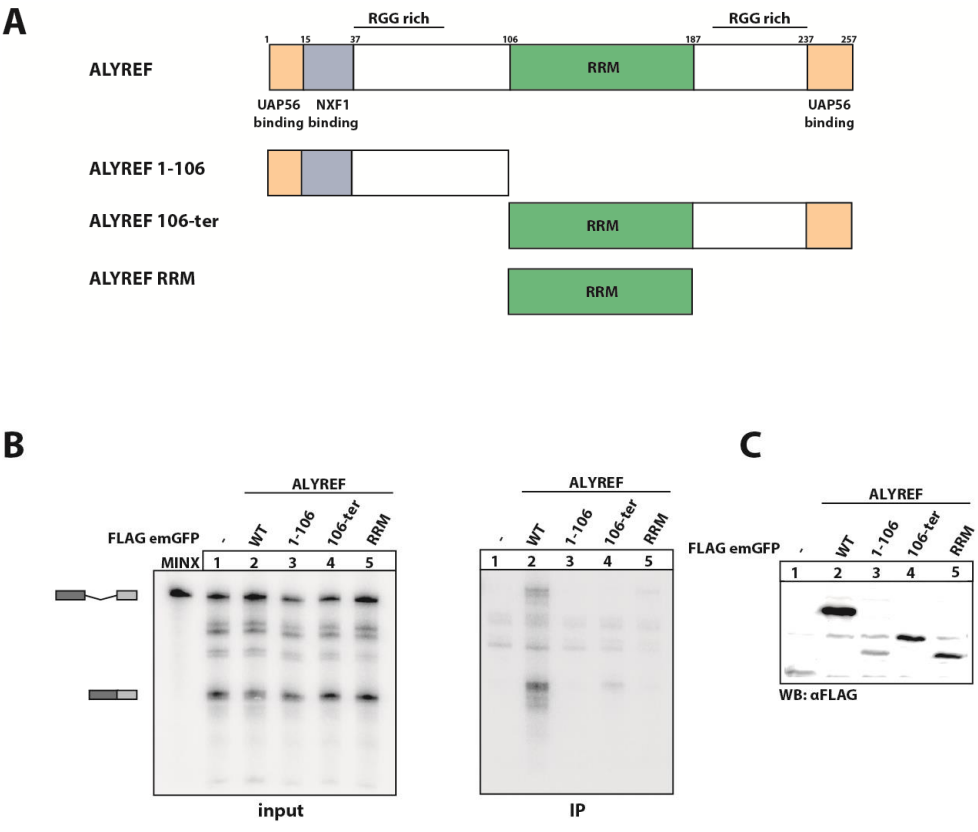


Figure 14. ALYREF predicted RNA-binding regions are not involved in precipitation of MINX mRNA. (A) Schematic representation of ALYREF mutants used in the study. **(B)** *In vitro* splicing reactions of MINX RNA supplemented with extracts expressing FLAG-tagged ALYREF truncation mutants (as in Figure 3). Schematic representations of splicing and digestions products are depicted on the left side of the autoradiograph. **(C)** Expression of

FLAG-tagged proteins in HEK 293 extract used in panel B was determined by immunoblot analysis using a FLAG antibody.

To understand the recruitment of ALYREF to RNA, all mutants lacking potential RNA-binding regions were subjected to the splicing assay (Figure 14B).

FLAG-tagged ALYREF efficiently precipitated spliced RNA (Figure 14B, lane 2). However, none of the deletion mutants precipitated spliced RNA, suggesting that neither the RGG-rich regions nor the RRM is able to bind RNA independently of other ALYREF domains (Figure 14 B, lane 3-5).

In order to confirm if these truncation mutants can fulfil their function in a specific cellular compartment, localization studies of FLAG emGFP fusion constructs were performed. As reported previously, full length ALYREF localizes to nuclear speckles (Rodrigues et al, 2001; Wichmann et al, 1999; Zhou et al, 2000). This feature was confirmed by a colocalization experiment with mCherry SC35 (Figure 15A).

In contrast to the full length ALYREF, none of the mutants retained localization to the nuclear speckles. ALYREF 1-106 localized to the nucleus. ALYREF 106-ter was predominantly localized to the nucleus, however some fluorescent signal was present in the cytoplasm. ALYREF RRM showed uniform localization throughout the cell (Figure 15B).

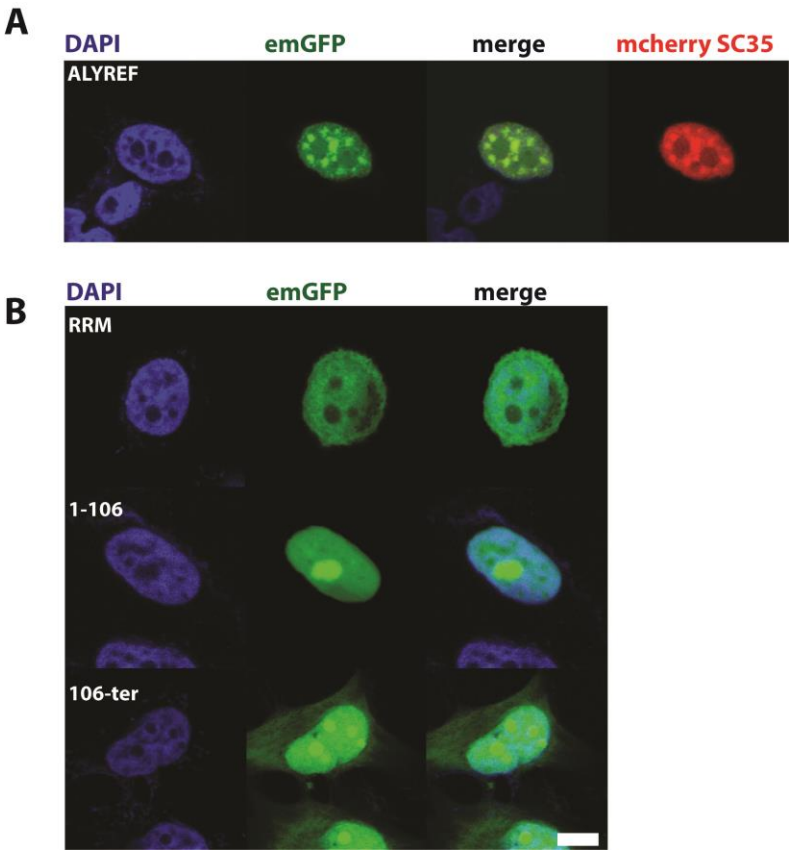


Figure 15. Localization of ALYREF and potential RNA-binding domains of ALYREF in the cell. (A) Colocalization of FLAG emGFP ALYREF with mCherry SC35 in HeLa cells. DAPI was used to stain nuclear DNA. Scale bar, 5 µm. **(B)** Localization of FLAG emGFP ALYREF truncation mutants in HeLa cells. DAPI was used to stain nuclear DNA.

ALYREF interacts with UAP56 through N-terminal and C-terminal sequences

To further investigate whether ALYREF requires interaction with UAP56 in order to be recruited to the mRNA during splicing, I used ALYREF mutants, lacking either the N-terminal or the C-terminal UAP56-binding sequences (Figure 16A). The importance of these sequences for a direct ALYREF:UAP56 binding was reported before (Hautbergue et al, 2009). Co-immunoprecipitation experiments were performed. Cell lysates expressing FLAG-tag ALYREF wild type or truncation

mutants were lysed and FLAG-immunoprecipitated. Co-immunoprecipitated UAP56 was detected using antibody raised against N-terminal peptide of UAP56. To avoid the possibility that interactions were mediated through RNA, lysates were treated with RNase A.

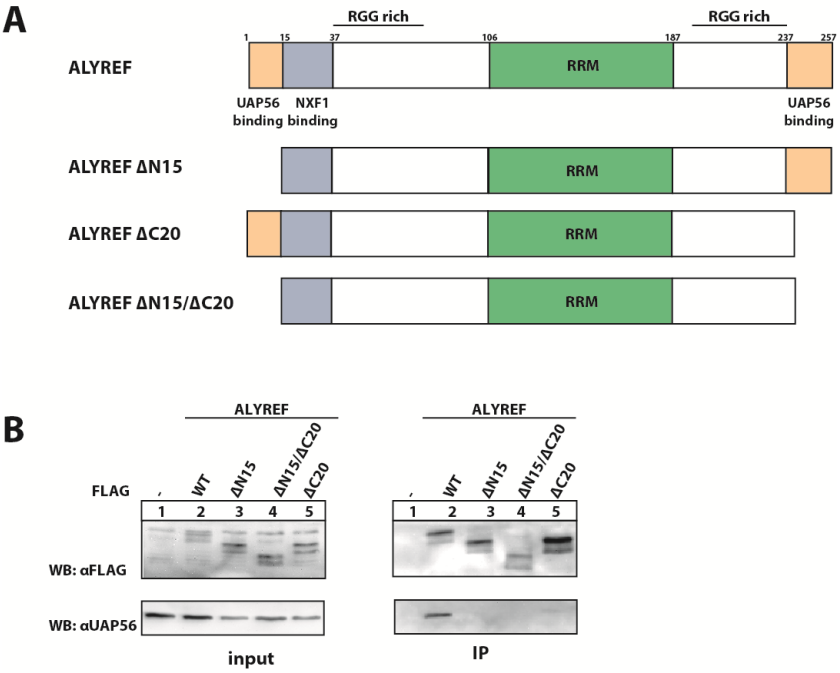


Figure 16. UAP56-binding domains of ALYREF are required for binding to UAP56. (A) Schematic representation of ALYREF mutants used in the study. **(B)** FLAG-immunoprecipitations of ALYREF mutants in RNase A-treated HeLa cell extracts. Co-immunoprecipitated UAP56 was detected by immunoblotting using a UAP56-specific antibody. 5 % of cell extracts were loaded as input. Extract expressing unfused FLAG-tag served a negative control.

In line with the previous reports, no UAP56 interaction was observed with the ALYREF mutant lacking both N- and C-terminal UAP56 binding sites. Moreover, deletion of N-terminal or C-terminal UAP56-binding motifs in ALYREF was sufficient to abolish interaction with UAP56 (Figure 16B).

ALYREF mutants lacking UAP56-binding sequences were subjected to the splicing assay (Figure 17). Surprisingly, in comparison to the wild type ALYREF, the truncation mutants showed only a mild decrease in binding to spliced mRNA (Figure 17A, compare lane 2 with lanes 3-5).

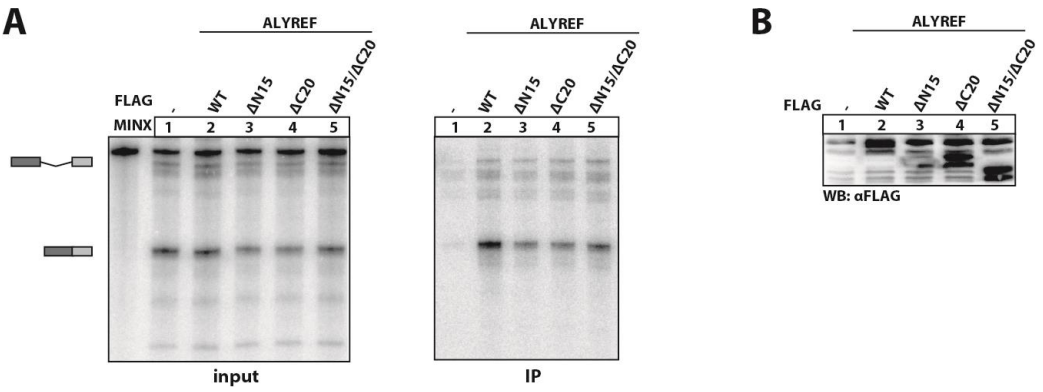


Figure17. ALYREF mutants lacking UAP56-binding precipitate spliced mRNA. (A) *In vitro* splicing reactions of MINX RNA supplemented with extracts expressing FLAG-tagged ALYREF truncation mutants (as in Figure 3). Schematic representations of splicing and digestions products are depicted on the left side of the autoradiograph. **(B)** Expression of FLAG-tagged proteins in HEK 293 extract used in panel B was determined by immunoblot analysis using a FLAG antibody.

Here I recapitulated the necessity of previously described ALYREF N- and C-terminal domains for UAP56 binding. Interestingly, mutants lacking UAP56-binding sites were still functional in the splicing assay (Figure 17A), suggesting that ALYREF recruitment to spliced mRNA is not dependent on interaction with UAP56.

UAP56 interaction is not required for ALYREF cap-specific deposition during splicing

UAP56 is an essential splicing factor, required for the first ATP-dependent step of splicing and spliceosome assembly at the branch point (Fleckner et al, 1997; Libri et

al, 2001). The recruitment of ALYREF to the mRNA by UAP56 was reported before (Dufu et al, 2010; Taniguchi & Ohno, 2008), although was not confirmed with experimental setting used in this study (Figure 17). Furthermore, I previously showed that both ALYREF and UAP56 binding to spliced mRNA is cap-dependent (Figure 5). Hence, I hypothesize that UAP56 might be required for the ALYREF recruitment to the cap, whereas specificity to the spliced RNA is mediated by a different region of ALYREF. Therefore, I subjected FLAG-tagged ALYREF ΔN15 (one of the ALYREF truncation mutants that did not precipitated UAP56 in Figure 16B) to the *in vitro* splicing assay in the presence of the excess of cap (Figure 18).

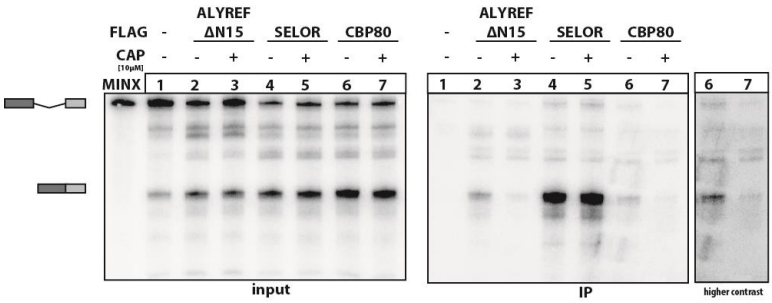


Figure 18. UAP56 interaction is not required for ALYREF cap-specific deposition during splicing. *In vitro* splicing reactions were performed as in figure 3. Splicing reaction of extract expressing FLAG-tagged ALYREF ΔN15 was supplemented with 10 μM of free cap (m^7GpppG) or not, mRNPs were immunoprecipitated and resolved on denaturing PAGE. FLAG-tagged SELOR serves as a negative control and CBP80 as positive control. Lanes 6 and 7 are represented separately using higher contrast for better visualization.

Surprisingly, FLAG-tagged ALYREF ΔN15 binding to spliced RNA was decreased in the presence of the cap excess (Figure 18, lane 2 and 3). This observation indicates that UAP56-binding is not required for ALYREF interaction with the cap.

ALYREF in vivo RNA-binding (CLIP experiments)

In order to elucidate positions where ALYERF binds to mRNA during export and mRNP formation, I utilized the CLIP assay. The CLIP assay was established for NXF1 (see section: NXF1 *in vivo* binding). Lysates expressing FLAG ALYREF were obtained cell line stably expressing protein of interest upon induction.

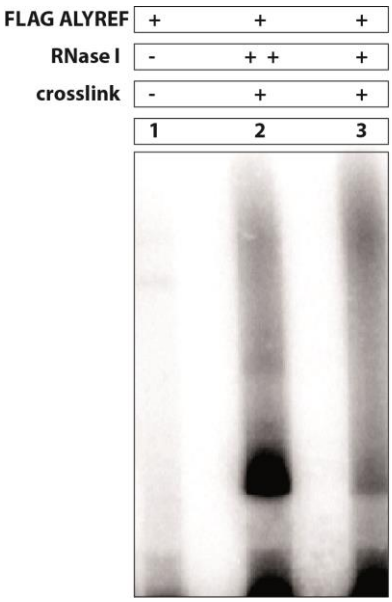


FIGURE 19. ALYREF can be *in vivo* crosslinked to mRNA. Lysates from Flp-In T-REx HEK-293 cell lines expressing FLAG ALYREF were crosslinked (+) or not (-) and digested with two RNAse I concentrations (++) for high RNase I and (+) for low RNase I concetration, crosslinked mRNA was labelled with PNK reaction.

After crosslinking, the mRNA was digested with different concentrations of RNase I. High RNase I concentration serves a control, the sharp band on the gel corresponds to the size of a protein. In the absence of UV-crosslinking, the band was not present on a gel (Figure 19, compare lane 1 and 2). In the low RNase I concentrations mRNA fragments were longer, what can be observed as an upshift

of mRNA fragment to the higher molecular weight (Figure 19, compare lane 2 and 3). This experiment showed that ALYREF not only stably precipitates mRNAs in splicing assays, but this binding can also be observed *in vivo.*

Identification of a novel RNA-binding motif

In this thesis I showed that ALYREF binds to the intact, capped, spliced mRNA. Both features of the mRNA, the presence of the cap and the EJC are required for ALYREF binding to mRNA. Furthermore, in contrast to previous findings, I showed that UAP56-biding is not required for ALYREF association with mRNAs.

To this end, the mechanism of ALYREF binding to spliced mRNA remains unclear. In the previous sections I ruled out the possibility of ALYREF recruitment to spliced mRNA through the interaction with UAP56 (Figure 18). Moreover, none of the known RNA-binding domains (RRM. RGG-rich regions) was alone mediating the interaction of ALYREF with mRNA (Figure 14).

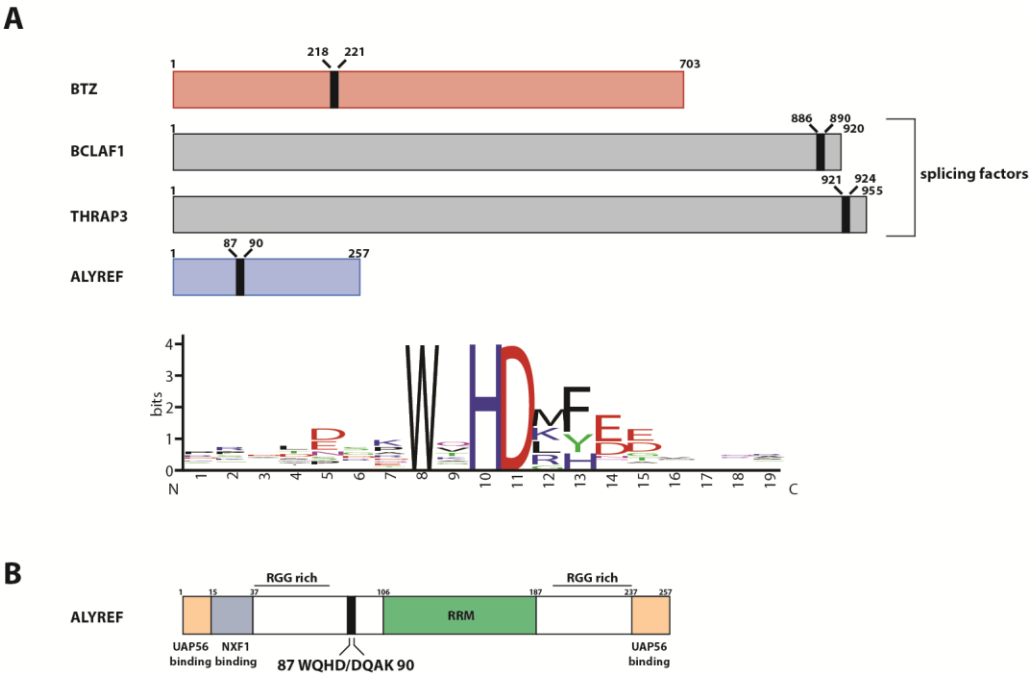


Figure 20. A novel, conserved motif is common among four RNA-binding proteins. (A) A short, conserved motif WxHD occurs in four RNA-binding proteins (BTZ, BCLAF1, THRAP3, and ALYREF). Weblogo diagram shows conservation of a motif in four proteins. **(B)** In ALYREF WQHD motif is placed in a low-complexity region. A point mutant 87 WQHD/ DQAK 90 was used in this study.

To understand which parts of ALYREF are responsible for its binding to the mRNA, I used the bioinformatic alignment of RNA-binding proteins (Kay Hoffman, personal communication). This alignment revealed a common, conserved motif – WxHD present only in THRAP3, BCLAF1, BTZ and ALYREF (Figure 20). THRAP3 and BCLAF1 are splicing factors (Lee et al, 2010; Lee et al, 2012), BTZ is a cytoplasmic component of EJC (Degot et al, 2004; Tange et al, 2005).

In ALYREF WxHD motif is located in a previously undescribed, unstructured region (87-90 aa) (Figure 20B). I utilized ALYREF mutant where WQHD was substituted with DQAK.

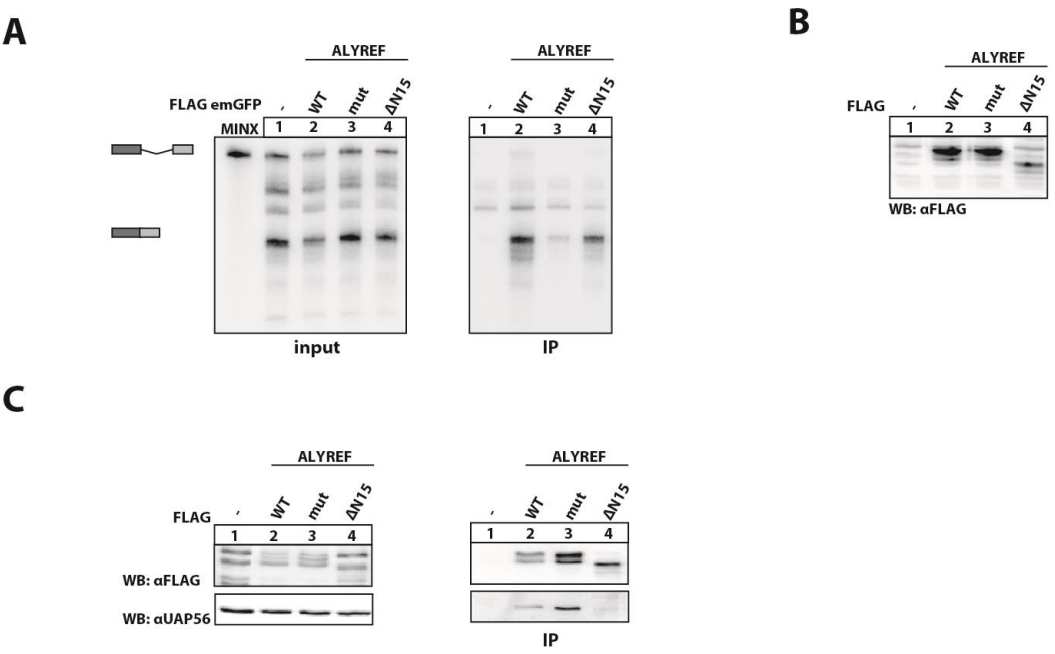


Figure 21. WQHD motif in ALYREF is required for efficient binding to spliced mRNA, not for ALYREF binding with UAP56. (A) *In vitro* splicing reactions of MINX substrate supplemented with extracts expressing FLAG-tagged ALYREF mutants (as in figure 3). Splicing products are schematically represented on the left side of a gel. **(B)** Expression of FLAG-tagged proteins in HEK 293 used in panel B determined by immunoblot analysis using a FLAG antibody. **(C)** FLAG-immunoprecipitations of ALYREF mutants in RNase A-treated HeLa cell extracts. Co-immunoprecipitated UAP56 was detected by immunoblotting using specific antibody. 5% cell extracts were loaded as input. Extract expressing unfused FLAG-tag served a negative control.

I subjected FLAG-tagged ALYREF 87 WQHD/DQAK 90 mutant to a splicing assay. In comparison to the wild type, mutant showed a reduced binding to spliced mRNA (Figure 21A, compare lane 2 and 3). In contrast, immunoprecipitation experiments showed that ALYREF 87 WQHD/DQAK 90 was still interacting with endogenous UAP56 to the same extend as the wild type (Figure 21C, lane 3). As shown in the previous section of this study FLAG-tagged ALYREF ΔN15 did not precipitate UAP56 (Figure 21C, lane 4).

In order to understand the cellular function of ALYREF 87 WQHD/DQAK 90 mutant, I performed the colocalization experiment with a known splicing factor – SC35 (Figure 22).

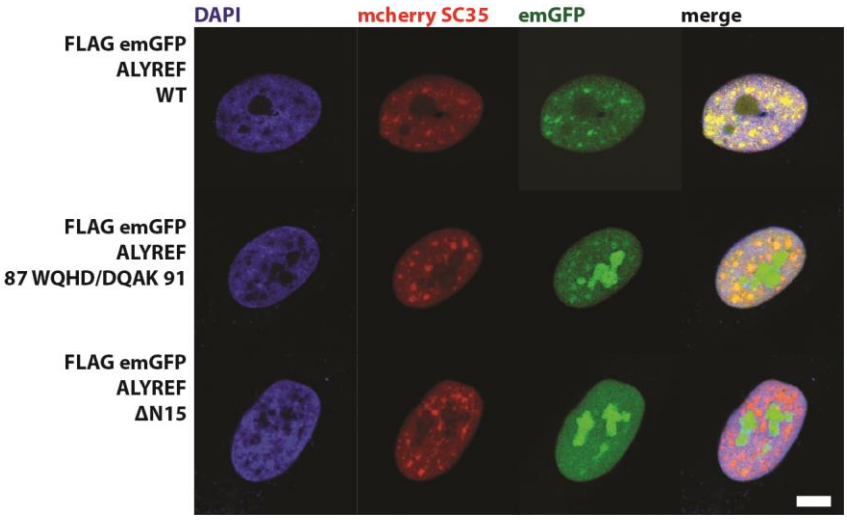


Figure 22. Colocalisation of ALYREF mutants with SC35. Localization of ALYREF mutant in HeLa cells. DAPI was used to stain nucleus. Scale bar, 5 µm.

In this experiment a fusion construct mCherry SC35 is localized to the nuclear speckles, as FLAG emGFP ALYREF. Nuclear speckles are defined, dynamic cellular regions where splicing components are stored (Girard et al, 2012). In contrast, FLAG emGFP ALYREF ΔN15 (lacking UAP56-interaction site) was localized to the nucleus, but not to the speckles. ALYREF 87 WQHD/DQAK 90 exhibited an intermediate phenotype, where only a portion of a protein maintained localization to the speckles (where the wild type ALYREF was localized).

These data suggest that interaction with UAP56, a spliceosome component, is required for a proper localization of ALYREF to the nuclear speckles. In the case of

The ALYREF 87 WQHD/DQAK 90 can be recruited to the nuclear speckles, probably due to the interaction with UAP56, during later steps of export, the protein dissociates from the mRNPs, what can be observed by the decreased binding to spliced mRNA in the splicing assay.

Monitoring mRNA export dynamics in the cell

In previous sections I examined how export factors are recruited to mRNA during mRNP assembly. In this section I established a technique enabling the visualization of mRNA export in cells.

One of the classical methods to visualize mRNA in the cell is fluorescent *in situ* hybridization (FISH) using oligo dT probe fused to the fluorophore (Molenaar et al, 2004; Politz et al, 1998; Politz et al, 1999) . In this thesis, cells were fixed and permeablized, afterwards the pre-hybridization buffer was applied. After washing, oligodT_{50} fused to Alexa 568 was added, afterwards cells were mounted and studied using confocal microscope.

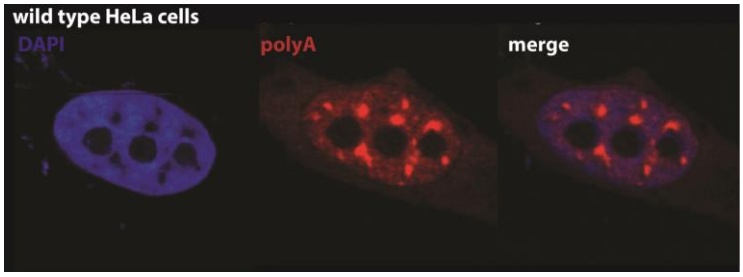


Figure 23. Localisation of polyadenylated mRNA in HeLa cells. FisH using oligodT_{50} fused with Alexa 568 probe. DAPI was used to stain nuclear DNA.

In this assay a signal is spread through the cell with few bright regions in the nucleus that represent transcription sites, where RNA accumulates (Figure 23). The nucleus was stained with DAPI.

Following single transcripts in the cell

FISH for polyadenylated mRNAs (Figure 23) can be used to observe only bulk polyadenylated mRNA in the cell. The specificity is too low to determine the localization of specific transcripts. To improve the specificity, recent studies on mRNA export dynamics focused on tracing single transcripts in living cells of a variety of models, such as: yeast (Bertrand et al, 1998), fruit fly (Jaramillo et al, 2008), mouse (Lionnet et al, 2011) and cell culture (Mor et al, 2010) .

I used a protocol that enabled me to visualize a single transcripts in the cell. Briefly, HEK-293 cell lines were stably expressing upon induction two reporter genes, either human β-globin gene (HBB) or green fluorescent protein (GFP). The transgene was incorporated into one specific loci within chromosome. FisH probes were complementary to a sequence within the 3′UTR, upstream of the open reading frame of the reporter genes. To enhance detection, the sequence complementary to the probe was inserted into the 3′UTR in several copies to form arrays. A schematic representation of arrays repetitions in 3′UTR of reporter genes are shown in Figure 24.

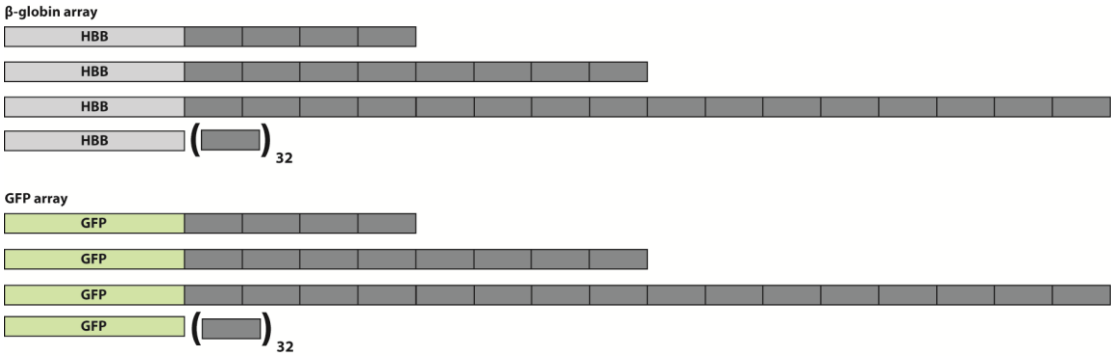


Figure 24. Schematic representation of reporter mRNAs used in the single-transcript FisH method. Grey box represents one repetition of a sequence recognized by the probe, an array.

The assay was tested on two cell lines with integrated β-globin arrays (HBB-16x array and HBB-32x array) (Figure 25). Without induction no reporter RNA was expressed, therefore no signal was observed in the cytoplasm (Figure 25). The pointed nuclear signal was arising from the chromosome locus, where the transgene was inserted. 24h after induction arrays were exported to the cytoplasm, which could be visualized by a uniformed punctuated signal in the cytoplasm. In the case of HBB-32x array more transcripts were retained in the nucleus in comparison to the HBB-16x array. In cells expressing HBB-16x array the puncta were evenly distributed throughout the cell. In both HBB-16x array and HBB-32x array a strong signal was detected from a single location in the nucleus, likely representing the single transcription site. For further functional experiments HBB-16x array was used, because of its favourable distribution within the cell.

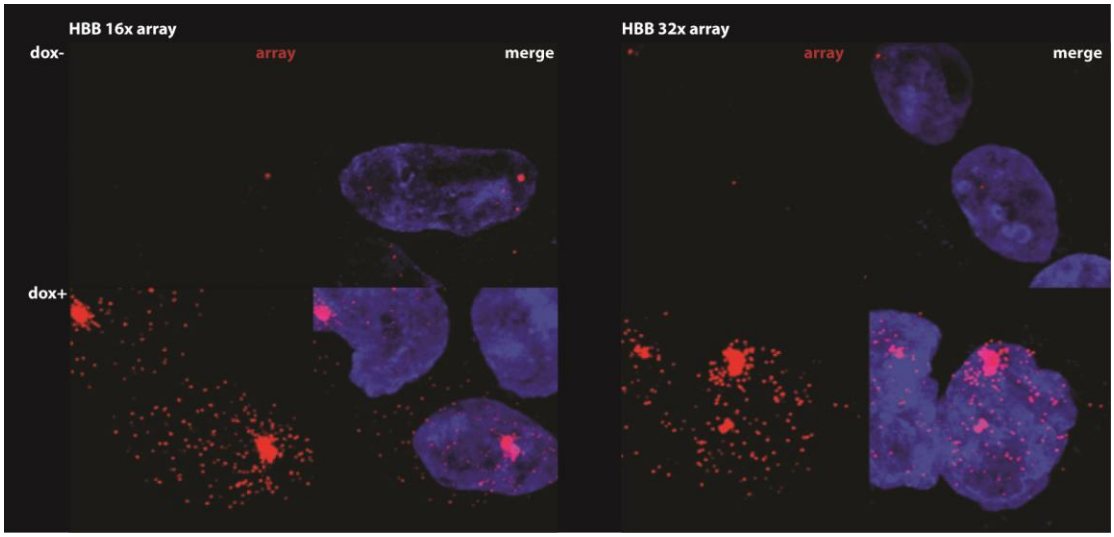


Figure 25. FlpIn T-REx HEK-293 with integrated reporter β-globin genes (HBB-16x array and HBB-32x array). Nucleus was stained with DAPI, reporter mRNA with probe against the array signal fused to the Alexa 568.

Following transcripts in the cell over time

The new method established in this thesis combines the specificity of a single transcript detection with the enhanced visualisation of multiple probe-tagging of the 3' UTR array repetition.

To this end, I established an assay in which reporter mRNA transcripts could be visualized in the cell, due to the presence 3 ′ UTR arrays. To understand how mRNA export functions over time, the chase experiment was performed.

Briefly, I utilized the HBB-16x array reporter cell line. Cells were seeded and the expression the reporter mRNA was induced. After specific time points (0h, 0,5h, 1h, 2h, 3h and 4h), cells were harvested and subjected them to FisH (Figure 26).

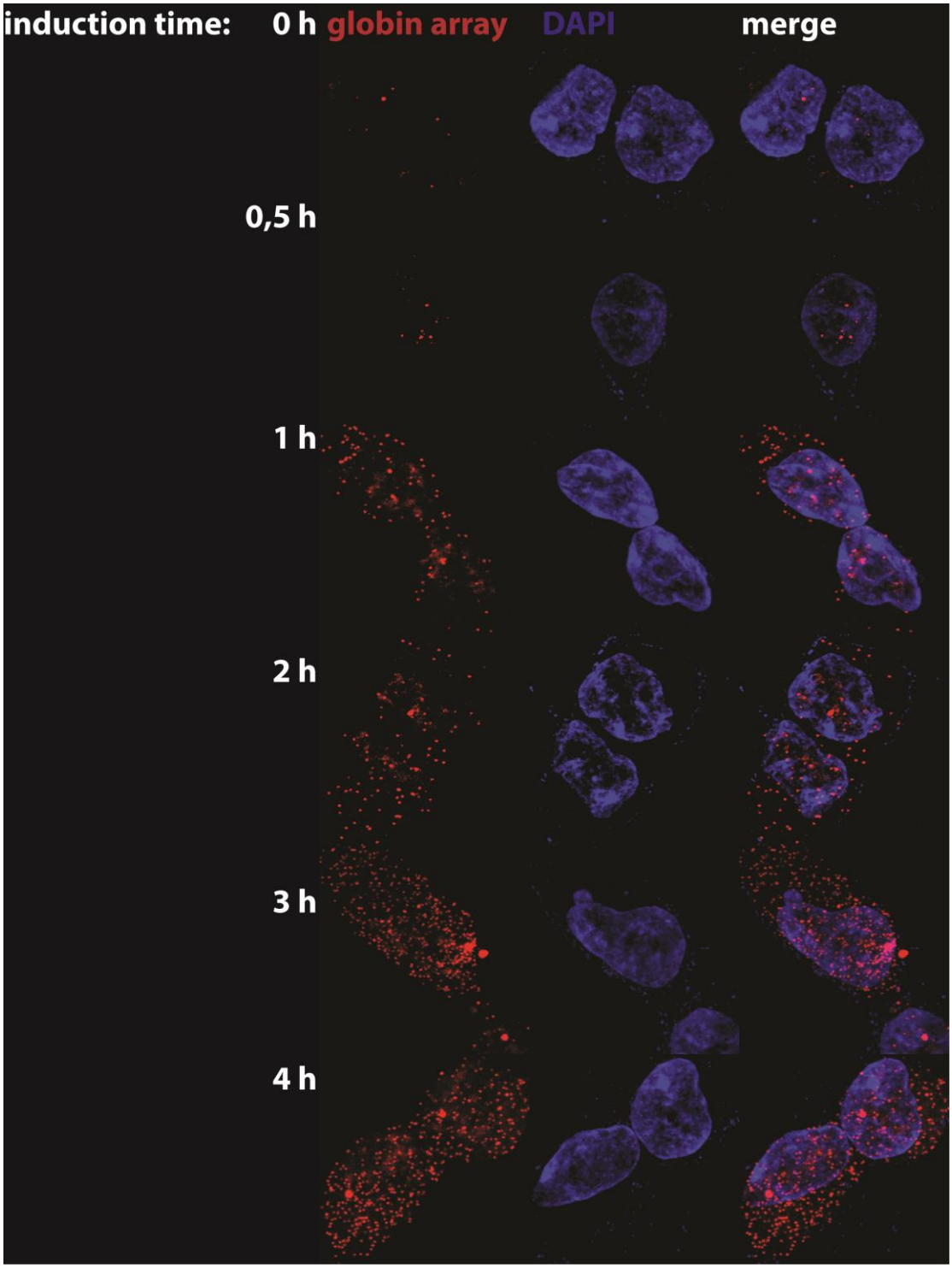


Figure 26. Expression of reporter mRNA (HBB-16x array) over time after induction. Nucleus was stained with DAPI, reporter mRNA with probe against the array signal fused to the Alexa 568.

In the chase experiment 30 min after induction of the transcript could be visualized in the nucleus (upper panel). After 60 min a few transcripts were exported to the cytoplasm (second panel). Furthermore, within the next 3 hours transcripts were evenly distributed throughout the cell (four bottom panels).

The chase experiment shows that searching for the single transcripts by FisH enables to monitor immediate changes in specific mRNA abundance in cells.

Following transcription and translation in the cell

In order to visualise translated protein from the reporter mRNA, the β-globin ORF was replaced by GFP ORF, to produce GFP arrays (Figure 24). The advantage of this reporter system its ability to simply track the functional product of mRNA export i.e. GFP. In the functional test of GFP-arrays (Figure 27) 24h after induction all tested GFP-arrays were successfully translated, which was monitored by green fluorescence (Figure 27).

In the case of the GFP-4x array the signal was evenly distributed throughout the cytoplasm and excluded from the nucleus, except for a few puncta corresponding to the insertion of the transgene into a chromosome (Figure 27, upper panel). For the GFP-8x array reporter cell line, after induction the transcripts were efficiently exported to the cytoplasm and single punctuated signals could be detected (Figure 27, middle panel). However, the abundance of the signal was lower in comparison to GFP-4x. In case of the GFP-16x array the majority of the transcripts was retained

in the nucleus (Figure 27, bottom panel), similar to the HBB 32x array (Figure 25). Surprisingly, the cytoplasmic amount of the GFP-16x array reporter mRNA was sufficient to obtain a relatively strong GFP signal.

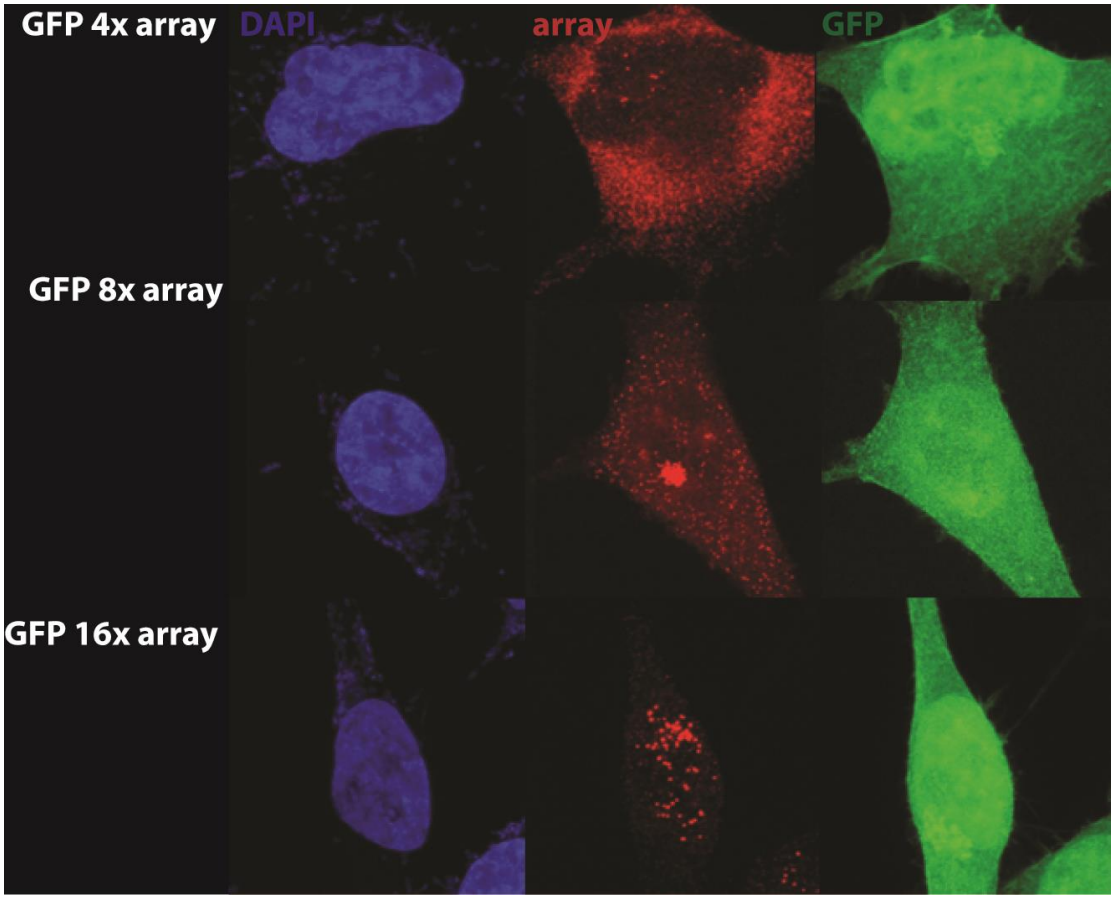


Figure 27. Expression of reporter mRNAs with various repletion of arrays fused to GFP 24h after induction. Nucleus was stained with DAPI, reporter mRNA with probe against the array signal fused to the Alexa 568.

In the further functional studies, I used the GFP-8x array, where the amount of inserted 3' UTR array was low, but high enough to detect single transcripts.

In case of GFP-8x array cell line, the low amount of array repetitions in the 3′UTR did not impair export, however provided sufficient number of binding site for the probe to enable detection.

mRNA retention upon NXF1 knockdown

To further characterize the GFP-8x array, I performed the knockdown of a general export receptor – NXF1. NXF1 knockdown was previously described to cause retention of polyadenylated mRNA in the nucleus in S2 cells (Herold et al, 2001) and mammalian cells (Hautbergue et al, 2008; Viphakone et al, 2012).

Briefly, the expression of the GFP-8x array was induced, at the same time two concentrations of siRNAs were used to knockdown endogenous NXF1. As a control a mock siRNA was used (Figure 28).

24 h after siRNA knockdown and simultaneous induction cells were fixed and FisH experiments were performed as described before. In control siRNA knockdown conditions, mRNA was evenly distributed throughout the cell (Figure 28, upper panel). In the case of NXF1 knockdown conditions, reporter mRNAs were retained in the nucleus (Figure 28, bottom panels). When the high concentration of siNXF1 (20 μM) were used, GFP signal decreased, suggesting that reduced amounts of reporter mRNA caused reduced translation of a functional protein.

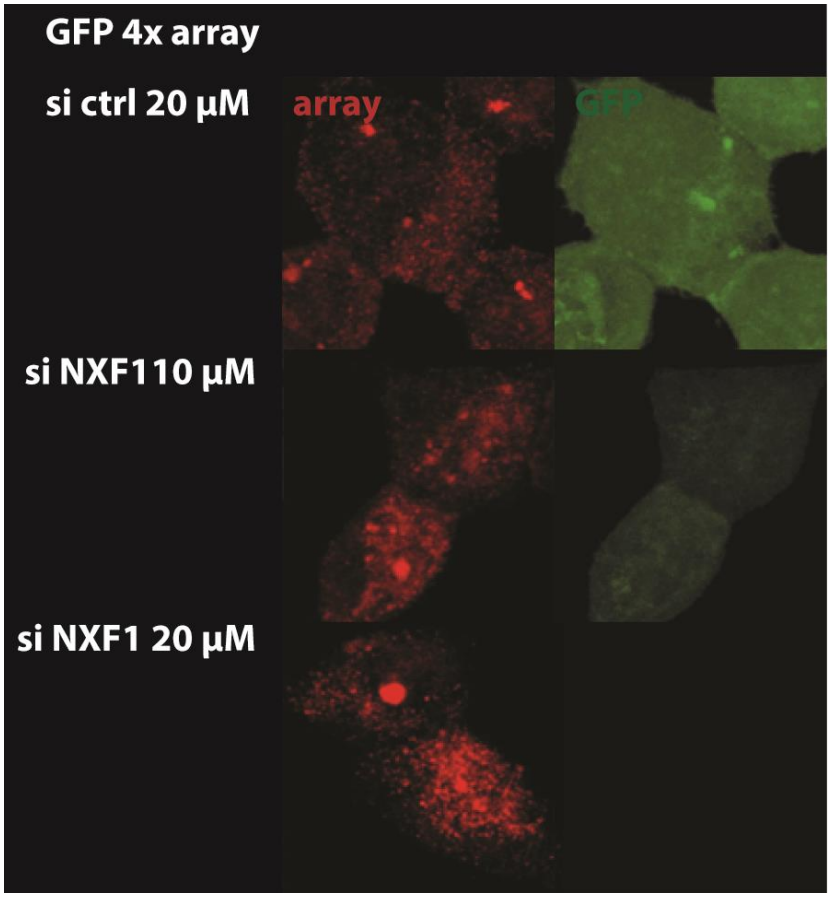


Figure 28. NXF1 knockdown results in nuclear retention of reporter mRNA. Reporter mRNA with probe against the array signal fused to the Alexa 568.

In this assay, I showed that the inducible reporter stable cell lines with 3′UTR arrays can be used to asses functional implications of depletion of export as it was shown for NXF1 (Figure 28).

Discussion

Export complex components differentially bind to mRNA during mRNP formation

This thesis presents the comprehensive analysis of known export factors and their *in vitro* binding to various mRNA targets. I find that only a core set of adaptor proteins stably interacts with spliced RNA. This set includes ALYREF, UAP56 and DDX39, which exhibit relatively strong binding to spliced mRNA (Figure 3). However, the recently described export adaptors (CIP29 and CHTOP) also precipitated spliced mRNA, although to a lesser extent (Figure 8). I did not observe any binding of the THO complex components to the MINX reporter mRNA (Figure 7).

Previous reports showed that THOC2 and THOC5 specifically bind to spliced mRNA *in vitro* (Masuda et al, 2005a), therefore the lack of interaction between mRNA and the THO complex components observed in this thesis was surprising. However, the genome-wide analysis of mRNAs regulated by the THO complex revealed that only a small proportion of mRNAs requires the THO complex for export (Rehwinkel, Herold et al. 2004), which is constistent with my results. Moreover, recent findings showed that THOC5 knockdown has only a mild effect on mRNA export (Chi et al, 2013; Tran et al, 2013). Thus, in contrast to yeast system, where the THO complex

is an essential component of export machinery, the mammalian THO complex potentially adopted a divergent function.

I found that the export adaptors CIP29 and CHTOP weakly interacted with spliced mRNA in comparison to ALYREF (Figure 8). This suggests that these export adaptors execute separate functions in mRNP assembly. Moreover, it was suggested that specific export adaptors bind to mRNA only during specific conditions (Hautbergue et al, 2009). Therefore, the core components (ALYREF, UAP56 and DDX39) bind stably with spliced mRNA, whereas the additional components (CIP29 and CHTOP) only temporarily associate with mRNPs to modulate the composition and thereby function of export complexes.

Next, I reexamined the role of the cap and the EJC in the recruitment of export factors. Initially, it was found that export factors preferentially associate with the 5' end of spliced RNAs (Cheng, Dufu et al. 2006). In a modified splicing assay, where the reporter mRNA was digested after splicing with RNase H I tested binding of the core export components (ALYREF, UAP56 and DDX39) to mRNA. (Figure 6). Due to the incomplete digestion of the reporter mRNA, it was feasible to compare binding of export factors to the intact mRNA and simultaneously to the cleaved products containing the 5' cap or the EJC binding site. Surprisingly, neither of the cleaved fragments (containing either the EJC or the cap) was precipitated as efficiently as the intact, uncleaved RNA. In conclusion, the recognition of export-competent mRNA is mediated by the mutual influence of at least two different binding sites.

The mRNA binding of the core export adaptors ALYREF and UAP56 can be outcompeted by the excess of the cap (Figure 5). This observation is consistent with a finding that binding of export adaptors to intronless mRNA is decreased in the presence of the cap excess (Nojima et al, 2007). However, the exact the mechanism of the interplay between the cap and the export adaptors remains to be elucidated.

NXF1 and spliced mRNPs do not form a stable complex

In Figure 3 all core export adaptors (e.g. ALYREF, UAP56 and DDX39) bind specifically to spliced mRNA. However, NXF1 does not interact with spliced RNA in an *in vitro* assay. This observation can be explained by the recently described self-inhibition activity of NXF1. Due to the intramolecular binding between RBD (RNA-binding domain) and NTF-2 domain, RNA binding activity of NXF1 is hindered. Accordingly, the RBD is only exposed when NXF1 binds to ALYREF and THOC5 (Viphakone et al, 2012). To impair the self-inhibition mode I utilized truncation mutants lacking the NTF-2 domain, the C-terminal inhibitory component (Figure 10). Surprisingly, none of the tested C-terminal truncation mutants can restore binding to mRNA.

Notably, all mutants, as well as the wild type NXF1 interact with CTE. CTE is a viral RNA sequence that promotes delivery of unspliced or partially spliced viral transcripts to the cytoplasm (Cullen, 1998; Grüter et al, 1998). The structure of a complex of NXF1 and CTE has revealed the molecular details of this interaction

(Teplova et al, 2011). Notably, the affinity of NXF1 to the stem loop structure of CTE is three orders of magnitude higher than for mRNA (Bachi et al, 2000; Hautbergue et al, 2008), suggesting that the interaction of NXF1 with cellular mRNAs is mediated by a mechanism different from the interaction between CTE and NXF1. In a cellular context the transient protein-protein interactions with export adaptor proteins might stabilize the interaction between general export receptor and mRNA. Moreover, the transient interaction between mRNA and NXF1 may facilitate the release of mRNA after translocation through the nuclear pore and prevent its re-import, ensuring the high turnover of export-competent complexes.

In cells, NXF1 acts as a heterodimer together with P15 (Katahira et al, 2002). Therefore, in this work the influence of P15 on NXF1 binding to CTE was assessed. Enhanced binding of NXF1 to CTE was observed when P15 was co-expressed (Figure 11). Moreover, P15 alone precipitated substantial amounts of CTE. Consistently, P15 was reported to be a crucial export factor for cellular mRNAs (Guzik et al, 2001) and required for the shuttling activity of NXF1 (Katahira et al, 2002). In contrast, it was shown that the activity of NXF1 in mRNA export does not depend on the interaction with P15 (Braun et al, 2002). Notably, P15 was implicated not only in mRNA export, but also in a protein-export pathway (Black et al, 2001). Thus, the exact function of P15 in enhancing NXF1 activity is controversial and remains to be elucidated further.

ALYREF interaction with mRNA is independent from UAP56

ALYREF contains two domains (RRM and two RRG-rich regions) that were previously described to bind to RNA in living cells. (Strein, Alleaume et al. 2014). However, no interaction of the isolated RRM or RGG-rich regions with RNA was detected in this thesis (Figure 14). It was previously reported that ALYREF is recruited to spliced mRNA upon interaction with UAP56 (Zhou, Luo et al. 2000, Luo, Zhou et al. 2001, Masuda, Das et al. 2005, Taniguchi and Ohno 2008). This interaction was described to be mediated through the N-terminal and the C-terminal UAP56-interaction motifs (Hautbergue and Hung 2008). In contrast to previous findings, all mutants lacking UAP56-binding motifs still retained the ability to associate with spliced RNAs, although the ALYREF:UAP56 interaction was abolished. This indicates that the interaction of ALYREF with mRNA is UAP56-independent, rising a possibility that ALYREF interaction with spliced mRNA is mediated through another undescribed motif.

A short, conserved motif is required for ALYREF binding to spliced mRNA

ALYREF was described to be an EJC component and thus its binding to mRNA was suggested to be dependent on splicing (Le Hir et al, 2000a; Le Hir et al, 2000c). Alternatively in yeast model, ALYREF homologue – Yra1 is recruited to the mRNA through the interaction with Sub2 (splicing factor, UAP56 homologue) (Strässer &

Hurt, 2001; Zenklusen et al, 2002). Similarly, in mammals ALYREF was found to be recruited to mRNA through a direct interaction with UAP56 (Luo et al, 2001; Zhou et al, 2000). However in this study, I show that ALYREF can be efficiently crosslinked to mRNA, therefore indicating that ALYREF associates with mRNPs through a close contact with mRNA in living cells (Figure 19).

Although several different models have been proposed in the literature, the exact mechanism of ALYREF deposition on spliced mRNA remains unknown. Recently, a short, conserved motif (WxHD) present in ALYREF and other proteins (BTZ, BCLAF1 and TRAP3) involved in mRNP formation has been identified (Kay Hofmann, personal communication). In ALYREF, the WQHD motif is located at position 87-91 in an unstructured, disordered region between the N-terminal RGG-like domain and the central RRM. Substitution of two amino acids in this motif leads to the reduction of binding to spliced RNA, although the interaction with UAP56 is not affected (Figure 21).

A similar WxHD motif was previously identified in the SELOR domain of BTZ. BTZ binds eIF4A3 with two contact surfaces. One of them is encoded by the WEHD motif. Mutations within this motif (W218D; HD220AK) abolish the interaction with eIF4A3 and prevent BTZ from binding to mRNA and to other EJC components (Gehring, Lamprinaki et al. 2009). The function of the motif in two additional proteins – BCLAF1 (Bcl-2-associated transcription factor 1, also referred to as BTF) and TRAP3 (Thyroid hormone associated receptor 3, also referred to as TRAP150)

has not been described yet. Both proteins were initially found as spliceosomal components (Merz et al, 2007). BCLAF1 and TRAP3 are SR-rich proteins that colocalize with EJC components in the cell (Varia et al, 2013). Moreover, both proteins were detected in the EJC-interactome (Singh et al, 2012). Furthermore, TRAP3 was identified as an NXF1 interaction partner and was reported to be involved in nuclear degradation of mRNA (Lee et al, 2010). Although little is known about BCLAF1 and TRAP3, the presented findings suggest that these two proteins are involved in mRNP assembly. I hypothesize that WxHD may represent a core sequence, which can promote binding to other mRNP components. The specificity of these interactions may be determined by surrounding sequences and allow proteins with a WxHD motif to bind to a number of similar interaction partners, what is reflected by similar functions in mRNP assembly of all WxHD-containing proteins.

Low complexity regions in mRNP assembly

A newly discovered motif WxHD present in ALYREF, BTZ, BCLAF1 and TRAP3 might potentially belong to the growing family of short linear motifs (SLiMs) also referred to as eukaryotic linear motifs (ELMs) (Davey et al, 2012; Neduva & Russell, 2006). In contrast to protein-protein interactions mediated through folded domains, SLiMs are required for low affinity interactions. SLiMs are short (3-10 amino acids) and present in disordered regions (also referred to as low-complexity regions, LC),

lacking tertiary structures. The evolutionary plasticity of these motifs is high due to the absence of structural constrains to maintain correct folding (Jonas & Izaurralde, 2013). Functionally, SLiMs either bind surfaces of the folded domains of their binding partner or intrinsically form complex mRNP granules (Davey et al, 2012; Tompa & Csermely, 2004).

mRNP granules are dynamic structures containing RBPs and their target mRNAs. mRNP granules were reported to be involved in mRNA degradation, translation repression, mRNA surveillance and RNA-mediated silencing (Eulalio et al, 2007; Parker & Sheth, 2007). mRNP granules exhibit a specific biophysical characteristics and can form liquid droplets or hydrogels. These structures are formed according to the principle of liquid-liquid phase separation. mRNP granules formation requires low-affinity, multivalent interactions, that are provided with SLiMs (Han et al, 2012; Jonas & Izaurralde, 2013; Kato et al, 2012).

The formation of mRNP granules was recently studied *in vitro. In vitro* reconstituted RNP-like aggregates precipitated by biotinylated isoxasole contain various proteins that have been identified by a mass-spectrometry analysis. This list of proteins associated with mRNP granules contains also proteins with WxHD motifs (ALYREF, TRAP3 and BTZ) (Han, Kato et al. 2012, Kwon, Xiang et al. 2014). Thus, the cooperative binding of proteins with WxHD motifs within LC regions may contribute to the coordinated assembly of mRNP aggregates, which, for example, may display a large number of export signals for increased export efficiency.

mRNP granules are complex structures where mRNA and RBPs, due to the presence of LC regions and short, conserved motifs, form together a functional structure. Mutations in any of mRNP granules components might be deleterious to the cell. The most common cause of familial amyothropic lateral sclerosis (ALS) and familial frontotemporal degeneration (FD) is an increased amount of GGGGCC repeats in intron 1 of C9orf72 (DeJesus-Hernandez et al, 2011; Renton et al, 2011). The exceeding number of hexanucleotide expansions leads to the sequestration of various RBPs on $(GGGGCC)_n$ RNA and formation of toxic mRNP granules (also referred to as RNA foci) (Ash et al, 2013; Mori et al, 2013). Recently, the analysis of GGGGCC repeats revealed that ALYREF colocalized with toxic RNA foci in cerebral granule cells and motor neurons from patients suffering from ALS and FD (Cooper-Knock et al, 2014). Thus, suggesting a potential role of ALYREF in the formation of mRNP granules and the deleterious effect of ALYREF retention on malfunctioning mRNAs.

The role of WHXD motif in mRNA export

It has been speculated that splicing and deposition of EJCs enhance the recruitment of export components and thereby stimulate export of spliced mRNPs (Le Hir, Gatfield et al. 2001). The EJC protein BTZ is localized to the cytoplasm in a steady state and therefore may only join the EJC after export. Hence, the surface region of eIF4A3 that is used for interaction with BTZ is not occupied in the nucleus.

Therefore, it is feasible that ALYREF docks onto eIF4A3 as part of the EJC in the nucleus after splicing is completed. However, a stable interaction of ALYREF with the mRNA requires the cap, which provides an additional binding site, supported by the previous observation that export complexes are located between the cap and the first EJC on mRNAs (Cheng et al, 2006). Finally, the export-competent mRNPs are transported the nuclear pore by a transient interaction with NXF1. ALYREF and other export factors are disassembled from the mRNPs and the eIF4A3 binding site can interact with BTZ in the cytoplasm (Figure 29). Alternatively, an EJC-independent recruitment of ALYREF to mature mRNAs may involve direct binding of the WxHD motif to mRNA. In either case, it will be interesting to find out how the simultaneous interaction of ALYREF with two distinct regions of the mRNA is established.

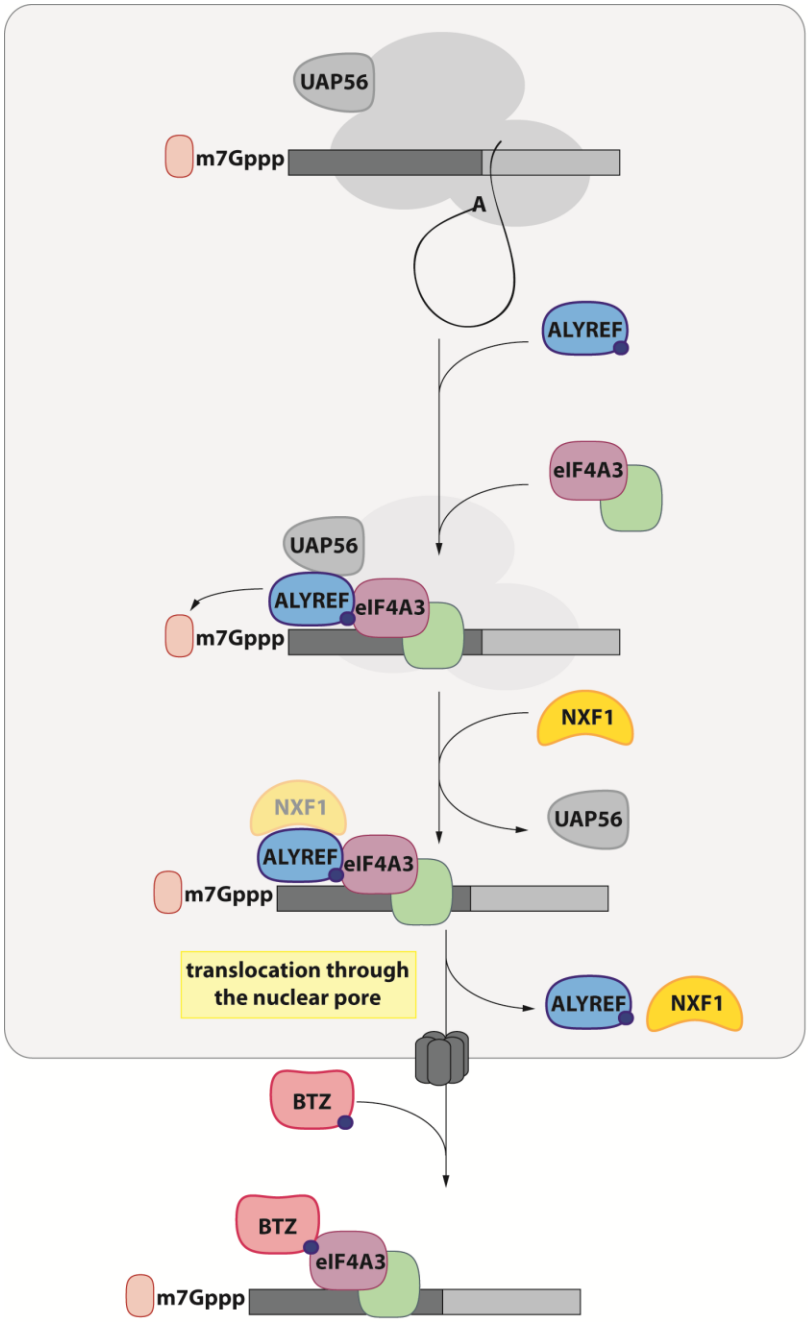


Figure 29. Model of mammalian export system.

Dynamics of mRNA export

One of the aims of this thesis was to establish a technique to visualize phenotype resulting from depletion of export components (Figures 24-28). Reporter mRNAs with 3′ arrays combine the specificity of a single-transcript tracing technique with

enhanced detection of multiple copies of the sequence recognized by the probe. The specificity and functionality of the method was confirmed by visualizing the nuclear retention of transcripts after NXF1-depletion (Figure 28).

Initially, to track dynamics of mRNA, transcripts in the cell were visualized using the MS2 system. In this system cells express reporter mRNA with MS2 binding sites and a fusion protein of MS2-capsid protein (MCP) (that can recognize MS2-binding site) with florescent protein (for detection purpose) and nuclear localization signal (Bertrand et al, 1998). The nuclear localisation signal on a fusion protein that recognizes MS2-binding sites prevents the unspecific florescent signal from MCP not bound to MS2-binding site in the cytoplasm.

In contrast to the MS2 system, where mRNA folds into a hairpin which serves as a recognition site for MCP, the reporter mRNA used in this thesis does not contain specific sequences that form additional secondary structures. Moreover, in this study no additional proteins are bound to the reporter mRNA, unlike in the MS2 system.

In this thesis I present a new technique to assess the localization of specific transcripts. However, in mammalian export system depletion of export adaptors (e.g. ALYREF, UAP56 and DDX39) does not lead to a global nuclear retention of polyadenylated transcripts (data not shown). Therefore, this method was not used for functional studies. Recent findings suggested that specific export adaptors only

bind to specific mRNAs, this can be detected by the high throughput analysis (e.g. CLIP) that identifies specific crosslink sites of factors to the mRNA and was applied in this thesis. Novel detection method presented in this thesis might be potentially used to validate information obtained from the CLIP assay to show how different mRNAs are exported upon different biological conditions.

Conclusions and Perspectives

Mammalian export is a crucial step in the mRNP life cycle. In this work, I examined the requirements of export adaptors and various RNA substrates for mRNP assembly. I found that ALYREF binds to spliced mRNA independently of UAP56 binding. A short motif (WQHD) within the low-complexity region of ALYREF is required for its interaction with mRNA. The exact mechanism of ALYREF binding to mRNA will need to be further examined in the future. Does the low complexity region of ALYREF directly interact with a specific site on mRNA, or is the interaction mediated by another factor? Based on the initial experiments presented in this thesis, it appears that *in vivo* crosslinking analysis could provide answers to this question.

Furthermore, the role of the THO complex remains elusive. There is growing evidence that these proteins mediate export of specific mRNAs rather than influence general export of mRNA. Therefore, the function of the THO complex in export of specific mRNAs needs to be assessed.

Only fully processed, mature mRNPs can be exported to the cytoplasm and translated. My results further corroborate the accumulated evidence that assembly of export adaptors requires both the cap and the EJC to be present, and may therefore serve as the last nuclear step controlling the correct assembly of RBPs on

mRNA. Therefore, mRNA export might be an additional quality control process that enables a quick release and retention of mRNAs, as a response to changing environmental conditions.

Materials and Methods

Materials and methods used in this study were prepared according to standard protocols (Sambrook et al, 2006). *In vitro* splicing assay and CLIP assay were performed accordingly to the published protocols (Huppertz et al, 2014; Steckelberg & Gehring, 2014).

Materials

Bacterial strains

For transformation purposes, XL1-blue *E.coli* (Stratagene) were used. Cells were grown in LB medium (10 g/l bacto-trypton, 10 g/l NaCl, 5 g/l yeast extract) or plates (LB medium supplemented 15 g/l agar) with the appropriate antibiotic.

Plasmids

Plasmid constructs were kindly provided by members of the Gehring Laboratory. Vectors used in this study: pCI-FLAG, pCI-FLAG-emGFP, pCI-V5, pCI-Venus, pcDNA5.

Tissue culture maintenance

The following cell-line strains were used during the project:

a) HEK-293T (ATTC) – human embryonic kidney transformed with SV40 which expresses the large T antigen.

b) Flp-InT-Rex 293 (Lite Technologies) – cells contain a single stably integrated tetracycline-inducible expression cassette

b) HeLa (Clonetech) – human cervical epithelial carcinoma.

Growth media for tissue culture

The cells were propagated in Dulbecco's Modified Eagle Media (DMEM) with GlutaMAX (Life Technologies), supplemented with 10% fetal calf serum (Life Technologies), 1% penicillin-streptomycin (Life Technologies).

All cells were maintained and transfected at 37 °C and 5% CO_2

Antibodies

antibody	**origin**	**dilution**
FLAG	Sigma-Aldrich	1:3000
V5	QED Bioscience	1:3000

Rabbit-HRP	Jackson ImmunoResearch	1:5000
UAP56[1]	Thermoscientific	1:3000

[1] raised against a peptide corresponding to 1-304 amino acid sequence within human UAP56

siRNA sequences

siRNA name	**siRNA sequence 5'-3'**
luciferase	AACGTACGCGGAATACTTCGA
NXF1	GGAUAUCUAUCAUCAUCAATT[1]

[1] the siRNA sequence from: (Katahira et al, 2009)

Buffers

All buffers which were required to be sterile were sterilised by autoclaving or by filtering through a 0.2 µm filter.

Transfection

2x BBS

50 mM BES

1.5 mM Na2HPO4

280 mM NaCl

pH 6.96

SDS page

Stacking gel (5ml)

0,7 ml 30%acrylamide/bisacrylamide (37,5:1)

2,975 ml water

1,25 ml 0,5 M Tris-HCl, pH 6,8

50 µl 10% (w/v) SDS

50 µl 10% (w/v) APS

5 µl TEMED

6x SDS-PAGE loading buffer

0.35 M Tris/HCl pH 6.8

10.28 % (w/v) SDS

36 % (v/v) glycerol

0,03% (w/v) bromophenolblue

Resolving gel 12% (10 ml)

4 ml 30% acrylamide/bisacrylamide (37,5:1)

3,2 ml water

1,25 ml 1,5 M Tris-HCl, pH 8,8

100 µl 10% (w/v) SDS

50 µl 10% (w/v) APS (Roth)

5 µl TEMED (Roth)

Running Buffer

25 mM Tris-HCl

200 mM Glycine

0,1% (w/v) SDS

Western blotting

Transfer buffer

25 mM Tris

192 mM glycine

10x TBS

50 mM Tris pH 7.4

1,5 M NaCl

20 % MeOH

1x TBS-Tween

1x TBS

0.2 % (v/v) Tween-20

Blocking solution

1x TBST

5 % (w/v) low fat milk powder

Immunoprecipitation

Lysis buffer

50 mM Tris [pH 7.2]

150 mM NaCl (Roth)

0.5% (v/v) Tween (Sigma)

50 µg/ml RNase A (Sigma)

1:200 protease inhibitor (Roche)

Wash buffer

50 mM Tris, pH 7.2

150 mM NaCl (Roth)

0.5% (v/v) Tween (Sigma)

CLIP

PNK Buffer

20 mM Tris-HCl, pH 7.4

10 mM $MgCl_2$ (Roth)

0.2% (v/v) Tween-20 (Sigma)

PNK mix

0.4 µl PNK (NEB)

0.8 µl γ^{32}P-ATP

0.8 µl PNK buffer (forward) (NEB)

6 µl H_2O

High salt wash

50 mM Tris-HCl, pH 7.4

1 M NaCl

1 mM EDTA

1% Igepal CA-630 (Sigma)

0.1% (w/v) SDS

0.5% (w/v) Na deoxycholate (Roth)

Lysis Buffer

50 mM Tris-HCl, pH 7.4

100 mM NaCl

1% Igepal CA-630 (Sigma)

0.1% (w/v) SDS

0.5% Na-deoxycholate (Roth)

Immunoflurescence and FiSH

Fixing solution

3,7 % (v/v) formaldehyde

1x PBS (Life technologies)

Permeabilising solution

0.5% (v/v) Triton

1x PBS

Hybridisation buffer

20 ml formamide (Roth)

10 ml (10 mg/ml tRNA (Sigma)

10 ml 20X SSC

20 ml dextran sulphate (Sigma)

37 ml H_2O

20x SSC

2M NaCl (Roth)

300 mM sodium citrate

Dehydrate (Roth)

Nuclei staining solution

1 µg/ml DAPI (Roth)

1x PBS (Life Technologies)

Mounting medium

6 g glycerin

2,4 g Mowiol 4-88 (Roth)

6 ml H_20

12,0 ml 0,2 M Tris-HCl, pH 8,5

625 mg DABCO (Roth)

***In vitro* splicing**

Buffer E

20 mM Hepes-KOH, pH 7,9

100 mM KCl

0,2 mM EDTA

10% (v/v) glycerol

1 mM DTT

EJC buffer

20 mM Hepes-KOH, pH 7,9

200 mM NaCl (Roth)

2 mM $MgCl_2$ (Roth)

0,2% (v/v) Triton X-100 (Sigma)

0,1% (v/v) Nonident P40 (Sigma)

0,05% (v/v) Na-deoxycolate (Roth)

1x TBE

89 mM Tris

89 mM boric acid

2 mM EDTA

RNA loading solution

80% formamide (Roth)

2 mM EDTA

Bromophenol blue

Xylencyanol

In vitro transcription reaction

2,5 µl α^{32}P-GTP (800 Ci/nmol)

20 u SP6 RNA polymerase (NEB)

1 µl trasncription buffer (NEB)

0,5 mM cap analogue (Promega)

1 mM DTT (Roth)

0,5 mM ATP

0,5 mM CTP

0,5 mM UTP

0,05 mM GTP

1 µl RNase inhibitor (Invitrogene)

In vitro splicing reaction

20% HeLa nuclear extract (CIBiotech)

0,2 mM ATP

40 mM creatine phosphate (Sigma)

5,4 mM $MgCl_2$ (Roth)

4,8% polyvinyl alcohol (Sigma)

44 mM HEPES, pH 7,3 (Roth)

RNase inhibitor (Invitrogene)

Denaturating PAGE gel mix

42% (w/v) urea (Roth)

10% Acrylamide (19:1) (Roth)

1x TBE

Glycogene blue

50% (v/v) glycogene (Roche)

Xylencyanol (Roth)

Deanturating gel

14 ml denaturing gel mix

140 µl 10% (w/v) APS (Roth)

20 µl TEMED (Roth)

Molecular Biology Kits

a) Small scale plasmid DNA purification: Zyppy Plasmid Miniprep Kit, Zymo Research.
b) Midi scale plasmid DNA purification: NucleoBond Xtra Midi, Macherey-Nagel
c) DNA extraction from agarose and PCR cleanup: NucleoSpin Gel and PCR clean-up

Methods

Plasmid transfections:

a) immunoprecipitation experiments (6cm dishes, 700 000 HeLa cells/dish seeded one day prior transfection):

pCI FLAG construct	1,5 µg
GFP (transfection control)	1 µg

b) *in vitro* splicing experiments (10 cm dishes, 1 200 000 HEK-293 cells/dish seeded one day prior transfection):

pCI FLAG construct	3 µg
GFP (transfection control)	2 µg

c) microscopy experiments (6 well plates, 200 000 HeLa cells/well seeded one day prior transfection):

pCI emGFP construct[1]	500 µg
pCI-mcherry construct[2]	1 µg
salmon sperm DNA	1 µg

[1] pCI-emGFP construct was used or mVenus (indicated in the Results section)

[2]in case of experiments with single fluorescent construct the amount of salmon sperm DNA was increased to 1,5 µg.

Cells were transfected using calcium phosphate method with BBS. The following concentrations of compounds were used, depending on the size of dish:

Size of cell culture dish	DNA and H_2O	2,5 M $CaCl_2$	2x BBS
6 well plate (2cm dish)	90 µl	10 µl	100 µl
6 cm dish	180 µl	20 µl	200 µl
10 cm dish	450 µl	50 µl	500 µl

Cell were seeded with appropriate concentrations one day prior the transfection. 4h prior the transfection, cell culture medium was exchanged and transfections were performed using appropriate transfection set up. The mix was incubated at RT for 15 min and applied to the cells. On the next day transfection efficiency was assessed.

siRNA knockdown

For NXF1 siRNA transfections, 200 000 HeLa cells were grown overnight in 6 wells plates (with coverslips) medium without antibiotics and transiently transfected with 150 pmol siRNA using Lipofectamine RNAiMAX (Life Technologies).

Immunoprecipitation

FLAG-complexes were immunoprecipitated from HeLa cell lysates. Cells were harvested 2 days after transfection in 400 µl of lysis buffer. Samples were kept overnight in -20°C and then centrifuged for 10 min at 10 000 x g at 4°C. 50 µl of

the lysate will serve as input. The rest of the sample was mixed with 12 µl of M2 anti-FLAG magnetic beads (Sigma) pre-washed in lysis buffer and incubated at 4°C overnight. On the next day the beads were washed three times (each of 5 minutes) with 600 µl wash buffer and complexes were eluted with 50 µl of SDS-sample buffer.

Western blot

Protein samples mixed with SDS-PAGE loading buffer were loaded on SDS-polyacrylamide gels. The proteins were transferred on Hybond-ECL nitrocellulose membrane (GE Healthcare) using Transfer buffer and the semi-dry system (Biorad). The transfer was performed for 90 minutes at 10 V. The membrane was incubated in blocking solution for 1h at RT. Next, the membrane was incubated overnight at 4°C with the primary antibody diluted in the blocking solution. After three washing steps (each of 10 minutes in TTBS), the membrane was incubated for 1h RT in the secondary antibody diluted in the blocking solution. Next, the membrane was washed three times (each of 10 minutes in TBS-tween) the membrane was incubated with ECL western blotting detection reagent (GE Healthcare) and developed with ImageQuant LAS 4000 (Thermoscientific).

Crosslinking and immunoprecipitation (CLIP)

Appropriate Flp-In T-Rex 293 cell lines expressing FLAG-tagged protein of interest were seeded (1 200 000 cells/10 cm dish) and induced for 3 days with 1 µg/ml of doxycycline. Next, medium was removed, cells were washed once with PBS (Life

Technologies) and crosslinked in 1 ml ice cold PBS once with 400 mJ in a UV-crosslinker (LTF Labortechnik) at 254 nm, harvested and frozen in -20°C as pellets. The non-crosslinked samples serve as controls.

On the next day, pellets were resuspended in 1 ml of lysis buffer (1:200 protease inhibitor added freshly). Samples were sonicated twice with 10 s bursts at 5 decibels (Branson sonifier). Samples were treated with different RNAse I (NEB) concentrations to obtain short RNA fragments (high RNase concentration – 1:50 dilution in lysis buffer) or long RNA fragments (low RNAse I concentration – 1:1000 dilution in lysis buffer). Each sample was treated with 2 µl of Turbo DNAse (Life Technologies) and 10 µl of appropriate RNAse dilution (NEB). Digestion was performed at 37 °C for 3 min at 1100 rpm, followed by 3 min incubation on ice.

After digestions, samples were centrifuged at maximum speed for 20 minutes. 40 µl of each sample was set aside as expression control for western blotting, the rest of the sample was incubated with 12 µl of M2 anti-FLAG magnetic beads (Sigma) pre-washed in lysis buffer for 2h at 4°C. After incubation time, beads were washed twice with high salt buffer and twice with PNK buffer. The last wash was removed from the beads that were incubated with 8 µl PNK mix at 37 °C for 5 min with shaking (1100 rpm). The PNK mix was removed and the samples were eluted in 20 µl of SDS loading buffer at 70 °C for 5 minutes. Samples were loaded on SDS PAGE gel and electrophoresis was performed. The gel was dried and exposed on a phosphoimager screen (Fuji BAS-MP IP) for 1h.

Immunofluorescence and fluorescent in situ hybridisation (FisH)

HeLa cells were seeded in 6 well plates and transfected accordingly. The day after transfection, the cells were transferred into new plates with coverslips. Two days after transfection the medium was removed, cells were washed with PBS (Life Technologies) and fixed for 15 min, the cells were incubated for 5 minutes on ice in permeabilising solution and washed with PBS (3 times, 5 min each).

Additional steps for FisH experiments

For FiSH experiments additional experimental procedures were performed. Coverslips were washed with 2x SSC buffer. Next the probe was applied (1:50 dilution in hybridisation buffer). I hybridisation was carried in 48°C for 1 hour in a hybridisation chamber. Cells were washed twice in 2x SSC buffer and twice with 1x SSC (each wash 5 min). After the last wash nuclei staining solution was applied on cells for 3 min. Cells were washed PBS and the cover slip were mounted on the microscope slide (Roth). Coverslips were kept overnight in 4 °C.

For immunofluorescence

After permeabilisation and washing cells were stained with DAPI and mounted on a glass slide as described above.

Microscopy and image acquisition

Images were obtained at room temperature on an Olympus FV1000 confocal microscope using a 60x UPlanApo (NA 1.35) objective. Assembly, contrast and

brightness adjustments of final image sets were performed using Photoshop CS4 (Adobe Systems, Inc.)

Generation of radiolabelled RNA substrates

DNA plasmid templates for *in vitro* transcription:

a) pGEM-4Z MINX
b) pGEM-4Z MINX Δi
c) pGEM-4Z AdmL ex2
d) pGEM-4Z CTE

To obtain template for *in vitro* transcription 8 µg of a pGEM-4Z vector with appropriate insert was digested with BamHI (NEB), DNA was purified on a gel and extracted using purification kit (Macherey-Nagel) and eluted according to the manufacturer's conditions. The transcription reaction was assembled (see section: Materials) and incubated for 40 min at 40°C. Next, Turbo DNase I (Life Technologies) digestion was performed at 37°C for 15 min. Afterwards, radioactively labelled, capped reporter RNA was extracted using Isol-RNA Lysis Reagent (5 Prime) and 1-bromo-3-chloropopane (Sigma) according to the manufacturer's protocol.

The radioactivity of the *in vitro* transcribed RNA was measured using scintillation counter (Beckman Coulter) and diluted to a working concentration of 10 000 cpm/µl.

Generation of FLAG-tagged protein extracts for in vitro splicing

293 cells were transfected in 10 cm dishes as described above. Cells were harvested two days after transfection in 300 µl buffer E and stored in -20 °C for at least 24h. Cell extracts were thawed on ice and sonicated (Branson sonifier) using 15 pulses, 30% intensity. Supernatants were collected for later use.

In vitro splicing, cap-excess and CTE binding assays

The *in vitro* splicing reaction mix (25 µl) was substituted with radiolabelled transcript (5 µl) and 20 µl of cell extracts (see above). Splicing reaction was performed for 90 min at 30°C. 8 µl of a reaction serves as input. The immunoprecipitation experiment was performed using 9 µl EZview Red ANTI-FLAG M2 beads (Sigma) prewashed in EJC buffer in 4°C for 2 hours. Next, samples were washed three times with EJC buffer and RNA was extracted using Isol-RNA Lysis Reagent (5 Prime) and 1-bromo-3-chloropopane (Sigma) according to the manufacturer's protocol. After overnight precipitation, RNA was resolved on denaturing gel. The gel was dried and exposed on a phosphoimager screen (Fuji BAS-MP IP) overnight.

In the cap excess splicing assay, *in vitro* splicing assay was modified accordingly. Splicing reactions were supplemented with 10 µM of m m^7GpppG RNA cap structure (NEB). After splicing, RNA-protein complexes were assembled, FLAG-tagged proteins were immunoprecipitated with associated RNA. RNA was resolved on denaturing gel and analysed.

In splicing assay followed by RNAse H digestion, *in vitro* splicing assay was modified accordingly. 10 µM of oligonucleotide was added (5′ CGG AAG AGA GTG 3′) after splicing reaction was completed. Digestion was performed for 15 min at 30 °C. After splicing and digestion, RNA-protein complexes were assembled, FLAG-tagged proteins were immunoprecipitated with associated RNA. RNA was resolved on denaturing gel and analysed.

CTE binding assays were performed using conditions established for *in vitro* splicing assay. Briefly, HeLa nuclear extract was supplemented with whole-cell extracts from HEK293 cells expressing FLAG-tagged proteins and radioactively labelled reporter CTE RNA. After CTE-containing complexes were assembled, FLAG-tagged proteins were immunoprecipitated together with associated CTE. RNA was resolved on denaturing gel and analysis.

References

Abruzzi KC, Lacadie S, Rosbash M (2004) Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *The EMBO journal* **23:** 2620-2631

Alexander RD, Innocente SA, Barrass JD, Beggs JD (2010) Splicing-dependent RNA polymerase pausing in yeast. *Mol Cell* **40:** 582-593

Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin W-LL, Dejesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW, Rademakers R, Boylan KB, Dickson DW, Petrucelli L (2013) Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77:** 639-646

Bachi A, Braun IC, Rodrigues JP, Panté N, Ribbeck K, von Kobbe C, Kutay U, Wilm M, Görlich D, Carmo-Fonseca M, Izaurralde E (2000) The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. *RNA (New York, NY)* **6:** 136-158

Ballut L, Marchadier B, Baguet A, Tomasetto C, Seraphin B, Le Hir H (2005) The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nat Struct Mol Biol* **12:** 861-869

Bear J, Tan W, Zolotukhin AS, Tabernero C, Hudson EA, Felber BK (1999) Identification of novel import and export signals of human TAP, the protein that binds to the constitutive transport element of the type D retrovirus mRNAs. *Molecular and Cellular Biology* **19:** 6306-6317

Bentley DL (2005) Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Current opinion in cell biology* **17:** 251-256

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM (1998) Localization of ASH1 mRNA particles in living yeast. *Molecular cell* **2:** 437-445

Black BE, Holaska JM, Levesque L, Ossareh-Nazari B, Gwizdek C, Dargemont C, Paschal BM (2001) Nxt1 Is Necessary for the Terminal Step of Crm1-Mediated Nuclear Export. *The Journal of Cell Biology* **152:** 141-156

Bono F, Ebert J, Lorentzen E, Conti E (2006) The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* **126:** 713-725

Braun IC, Herold A, Rode M, Conti E, Izaurralde E (2001) Overexpression of TAP/p15 heterodimers bypasses nuclear retention and stimulates nuclear mRNA export. *The Journal of biological chemistry* **276:** 20536-20543

Braun IC, Herold A, Rode M, Izaurralde E (2002) Nuclear export of mRNA by TAP/NXF1 requires two nucleoporin-binding sites but not p15. *Molecular and cellular biology* **22:** 5405-5418

Braun IC, Rohrbach E, Schmitt C, Izaurralde E (1999) TAP binds to the constitutive transport

element (CTE) through a novel RNA-binding motif that is sufficient to promote CTE-dependent RNA export from the nucleus. *The EMBO journal* **18:** 1953-1965

Buchwald G, Schüssler S, Basquin C, Le Hir H, Conti E (2013) Crystal structure of the human eIF4AIII-CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 8

Buratowski S (2009) Progression through the RNA polymerase II CTD cycle. *Molecular cell* **36:** 541-546

Carmody SR, Tran EJ, Apponi LH, Corbett AH, Wente SR (2010) The mitogen-activated protein kinase Slt2 regulates nuclear retention of non-heat shock mRNAs during heat shock-induced stress. *Molecular and cellular biology* **30:** 5168-5179

Carrillo Oesterreich F, Preibisch S, Neugebauer KM (2010) Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol Cell* **40:** 571-581

Chang CT, Hautbergue GM, Walsh MJ, Viphakone N, van Dijk TB, Philipsen S, Wilson SA (2013) Chtop is a component of the dynamic TREX mRNA export complex. *The EMBO journal* **32:** 473-486

Chavez S, Beilharz T, Rondon AG, Erdjument-Bromage H, Tempst P, Svejstrup JQ, Lithgow T, Aguilera A (2000) A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *The EMBO journal* **19:** 5824-5834

Chazal P-EE, Daguenet E, Wendling C, Ulryck N, Tomasetto C, Sargueil B, Le Hir H (2013) EJC core component MLN51 interacts with eIF3 and activates translation. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 5903-5908

Cheng H, Dufu K, Lee C-SS, Hsu JL, Dias A, Reed R (2006) Human mRNA export machinery recruited to the 5' end of mRNA. *Cell* **127:** 1389-1400

Chi B, Wang Q, Wu G, Tan M, Wang L, Shi M, Chang X, Cheng H (2013) Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic acids research* **41:** 1294-1306

Conti E, Izaurralde E (2001) Nucleocytoplasmic transport enters the atomic age. *Current opinion in cell biology* **13:** 310-319

Cooper-Knock J, Walsh MJ, Higginbottom A, Robin Highley J, Dickman MJ, Edbauer D, Ince PG, Wharton SB, Wilson SA, Kirby J, Hautbergue GM, Shaw PJ (2014) Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain : a journal of neurology* **137:** 2040-2051

Coyle JH, Bor Y-CC, Rekosh D, Hammarskjold M-LL (2011) The Tpr protein regulates export of mRNAs with retained introns that traffic through the Nxf1 pathway. *RNA (New York, NY)* **17:** 1344-1356

Cramer P (2004) RNA polymerase II structure: from core to functional complexes. *Curr Opin Genet Dev* **14:** 218-226

Cullen BR (1998) Retroviruses as model systems for the study of nuclear RNA export pathways. *Virology* **249:** 203-210

Davey NE, Van Roey K, Weatheritt RJ, Toedt G, Uyar B, Altenberg B, Budd A, Diella F, Dinkel H, Gibson TJ (2012) Attributes of short linear motifs. *Molecular bioSystems* **8:** 268-281

Davidson L, Muniz L, West S (2014) 3' end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes & development* **28:** 342-356

Degot S, Le Hir H, Alpy F, Kedinger V, Stoll I, Wendling C, Seraphin B, Rio M-CC, Tomasetto C (2004) Association of the breast cancer protein MLN51 with the exon junction complex via its speckle localizer and RNA binding module. *The Journal of biological chemistry* **279:** 33702-33715

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung G-YRY, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72:** 245-256

Dias AP, Dufu K, Lei H, Reed R (2010) A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat Commun* **1:** 97

Dreyfuss G, Kim VN, Kataoka N (2002) Messenger-RNA-binding proteins and the messages they carry. *Nature Reviews Molecular Cell ...*

Dufu K, Livingstone MJ, Seebacher J, Gygi SP, Wilson SA, Reed R (2010) ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes & development* **24:** 2043-2053

Edery I, Sonenberg N (1985) Cap-dependent RNA splicing in a HeLa nuclear extract. *Proceedings of the National Academy of Sciences* **82:** 7590-7594

Egloff S, Murphy S (2008) Cracking the RNA polymerase II CTD code. *Trends in genetics : TIG* **24:** 280-288

Eulalio A, Behm-Ansmant I, Izaurralde E (2007) P bodies: at the crossroads of post-transcriptional pathways. *Nature reviews Molecular cell biology* **8:** 9-22

Fleckner J, Zhang M, Valcarcel J, Green MR (1997) U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes & development* **11:** 1864-1872

Fribourg S, Conti E (2003) Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15. *EMBO reports* **4:** 699-703

Fukuda S, Wu DW, Stark K, Pelus LM (2002) Cloning and characterization of a proliferation-associated cytokine-inducible protein, CIP29. *Biochemical and biophysical research communications* **292:** 593-600

Gaillard H, Wellinger RE, Aguilera A (2007) A new connection of mRNP biogenesis and export with transcription-coupled repair. *Nucleic acids research* **35:** 3893-3906

Galy V, Gadal O, Fromont-Racine M, Romano A (2004) Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. *Cell*

Gatfield D, Izaurralde E (2002) REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. *The Journal of cell biology* **159:** 579-588

Gatfield D, Le Hir H, Schmitt C, Braun IC, Kocher T, Wilm M, Izaurralde E (2001) The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in Drosophila. *Current Biology* **11:** 1716-1721

Gehring NH, Lamprinaki S, Hentze MW, Kulozik AE (2009) The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. *PLoS biology* **7**

Girard C, Will C, Peng J, Makarov E, Kastner B, Lemm I, Urlaub H, Hartmuth K, Lührmann R (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nature communications* **3:** 994

Gomez-Gonzalez B, Garcia-Rubio M, Bermejo R, Gaillard H, Shirahige K, Marin A, Foiani M, Aguilera A (2011) Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *The EMBO journal* **30:** 3106-3119

Gonatopoulos-Pournatzis T, Cowling VH (2014) Cap-binding complex (CBC). *The Biochemical journal* **457:** 231-242

Grüter P, Tabernero C, von Kobbe C, Schmitt C, Saavedra C, Bachi A, Wilm M, Felber BK, Izaurralde E (1998) TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Molecular cell* **1:** 649-659

Gu B, Eick D, Bensaude O (2013) CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic acids research* **41:** 1591-1603

Guzik BW, Levesque L, Prasad S, Bor YC, Black BE, Paschal BM, Rekosh D, Hammarskjöld ML (2001) NXT1 (p15) is a crucial cellular cofactor in TAP-dependent export of intron-containing RNA in mammalian cells. *Molecular and cellular biology* **21:** 2545-2554

Han TW, Kato M, Xie S, Wu LC, Mirzaei H, Pei J, Chen M, Xie Y, Allen J, Xiao G, McKnight SL (2012) Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell* **149:** 768-779

Hautbergue GM, Hung M-LL, Walsh MJ, Snijders AP, Chang C-TT, Jones R, Ponting CP, Dickman MJ, Wilson SA (2009) UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Current biology : CB* **19:** 1918-1924

Hautbergue GM, Hung ML, Golovanov AP, Lian LY, Wilson SA (2008) Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc Natl Acad Sci U S A* **105:** 5154-5159

Herold A, Klymenko T, Izaurralde E (2001) NXF1/p15 heterodimers are essential for mRNA nuclear export in Drosophila. *RNA (New York, NY)* **7:** 1768-1780

Herold A, Suyama M, Rodrigues JP, Braun IC, Kutay U, Carmo-Fonseca M, Bork P, Izaurralde E (2000) TAP (NXF1) belongs to a multigene family of putative RNA export factors with a conserved modular architecture. *Molecular and cellular biology* **20:** 8996-9008

Hirose Y, Manley JL (1998) RNA polymerase II is an essential mRNA polyadenylation factor. *Nature*

Ho CK, Shuman S (1999) Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Molecular cell*

Huang Y, Gattoni R, Stévenin J, Steitz JA (2003) SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Molecular cell* **11:** 837-843

Huang Y, Steitz JA (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Molecular cell* **7:** 899-905

Huang Y, Steitz JA (2005) SRprises along a messenger's journey. *Mol Cell* **17:** 613-615

Huang Y, Yario TA, Steitz JA (2004) A molecular link between SR protein dephosphorylation and mRNA export. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 9666-9670

Huppertz I, Attig J, D'Ambrogio A, Easton LE, Sibley CR, Sugimoto Y, Tajnik M, König J, Ule J (2014) iCLIP: protein-RNA interactions at nucleotide resolution. *Methods (San Diego, Calif)* **65:** 274-287

Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj IW (1995) A cap-binding protein complex mediating U snRNA export. *Nature* **376:** 709-712

Izaurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78:** 657-668

Jaramillo AM, Weil TT, Goodhouse J, Gavis ER, Schupbach T (2008) The dynamics of fluorescently labeled endogenous gurken mRNA in Drosophila. *Journal of cell science* **121:** 887-894

Jensen T, Patricio K, McCarthy T, Rosbash M (2001a) A Block to mRNA Nuclear Export in< i> S. cerevisiae</i> Leads to Hyperadenylation of Transcripts that Accumulate at the Site of Transcription. *Molecular cell* **7:** 887-898

Jensen TH, Boulay J, Rosbash M, Libri D (2001b) The DECD box putative ATPase Sub2p is an early mRNA export factor. *Current biology : CB* **11:** 1711-1715

Jimeno S, Luna R, García-Rubio M (2006) Tho1, a novel hnRNP, and Sub2 provide alternative pathways for mRNP biogenesis in yeast THO mutants. *Molecular and cellular ...*

Jimeno S, Rondon AG, Luna R, Aguilera A (2002) The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *The EMBO journal* **21:** 3526-3535

Johnson SA, Kim H, Erickson B, Bentley DL (2011) The export factor Yra1 modulates mRNA 3' end processing. *Nat Struct Mol Biol* **18:** 1164-1171

Jonas S, Izaurralde E (2013) The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes & development* **27:** 2628-2641

Jurica MS, Moore MJ (2003) Pre-mRNA splicing: awash in a sea of proteins. *Molecular cell* **12:** 5-14

Kallehauge TB, Robert M-CC, Bertrand E, Jensen TH (2012) Nuclear retention prevents premature cytoplasmic appearance of mRNA. *Molecular cell* **48:** 145-152

Kang Y, Cullen BR (1999) The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. *Genes & development* **13:** 1126-1139

Kapadia F, Pryor A, Chang T-HH, Johnson LF (2006) Nuclear localization of poly(A)+ mRNA following siRNA reduction of expression of the mammalian RNA helicases UAP56 and URH49. *Gene* **384:** 37-44

Katahira J, Inoue H, Hurt E, Yoneda Y (2009) Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *The EMBO journal* **28:** 556-567

Katahira J, Straesser K, Saiwaki T, Yoneda Y, Hurt E (2002) Complex formation between Tap and p15 affects binding to FG-repeat nucleoporins and nucleocytoplasmic shuttling. *The Journal of biological chemistry* **277:** 9242-9246

Katahira J, Strasser K, Podtelejnikov A, Mann M, Jung JU, Hurt E (1999) The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *The EMBO journal* **18:** 2593-2609

Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, Grishin NV, Frantz DE, Schneider JW, Chen S, Li L, Sawaya MR, Eisenberg D, Tycko R, McKnight SL (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149:** 753-767

Kim H, Erickson B, Luo W, Seward D, Graber JH, Pollock DD, Megee PC, Bentley DL (2010) Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nat Struct Mol Biol* **17:** 1279-1286

Kistler AL, Guthrie C (2001) Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes & development*

Köhler A, Hurt E (2007) Exporting RNA from the nucleus to the cytoplasm. *Nature reviews Molecular cell biology* **8:** 761-773

Komarnitsky P, Cho EJ, Buratowski S (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & development* **14:** 2452-2460

Konarska MM, Padgett RA, Sharp PA (1984) Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **38:** 731-736

Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J (2011) iCLIP--transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. *Journal of visualized experiments : JoVE*

König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology* **17:** 909-915

Le Hir H, Gatfield D, Braun IC, Forler D, Izaurralde E (2001a) The protein Mago provides a link between splicing and mRNA localization. *EMBO reports* **2:** 1119-1124

Le Hir H, Gatfield D, Izaurralde E, Moore MJ (2001b) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *The EMBO journal* **20:** 4987-4997

Le Hir H, Izaurralde E, Maquat LE, Moore MJ (2000a) The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO journal* **19:** 6860-6869

Le Hir H, Izaurralde E, Maquat LE, Moore MJ (2000b) The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO journal* **19:** 6860-6869

Le Hir H, Moore MJ, Maquat LE (2000c) Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes & development* **14:** 1098-1108

Le Hir H, Moore MJ, Maquat LE (2000d) Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon–exon junctions. *Genes & development*

Leaw CL, Ren EC, Choong ML (2004) Hcc-1 is a novel component of the nuclear matrix with growth inhibitory function. *Cellular and Molecular Life Sciences*

Lee K-MM, Hsu I-Wa, Tarn W-YY (2010) TRAP150 activates pre-mRNA splicing and promotes nuclear mRNA degradation. *Nucleic acids research* **38:** 3340-3350

Lee YY, Yu YB, Gunawardena HP, Xie L, Chen X (2012) BCLAF1 is a radiation-induced H2AX-interacting partner involved in γH2AX-mediated regulation of apoptosis and DNA repair. *Cell death & disease* **3**

Libri D, Graziani N, Saguez C, Boulay J (2001) Multiple roles for the yeast SUB2/yUAP56 gene in splicing. *Genes & development* **15:** 36-41

Liker E, Fernandez E, Izaurralde E, Conti E (2000) The structure of the mRNA export factor TAP reveals a cis arrangement of a non-canonical RNP domain and an LRR domain. *The EMBO journal* **19:** 5587-5598

Linder P, Jankowsky E (2011) From unwinding to clamping - the DEAD box RNA helicase

family. *Nature reviews Molecular cell biology* **12:** 505-516

Lionnet T, Czaplinski K, Darzacq X, Shav-Tal Y, Wells AL, Chao JA, Park HY, de Turris V, Lopez-Jones M, Singer RH (2011) A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nature methods* **8:** 165-170

Long J, Caceres J (2009) The SR protein family of splicing factors: master regulators of gene expression. *Biochem J*

Luo ML, Zhou Z, Magni K, Christoforides C, Rappsilber J, Mann M, Reed R (2001) Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413:** 644-647

Masuda S, Das R, Cheng H, Hurt E, Dorman N, Reed R (2005a) Recruitment of the human TREX complex to mRNA during splicing. *Genes & development* **19:** 1512-1517

Masuda S, Das R, Cheng H, Hurt E, Dorman N, Reed R (2005b) Recruitment of the human TREX complex to mRNA during splicing. *Genes & development* **19:** 1512-1517

Mazza C, Ohno M, Segref A, Mattaj IW, Cusack S (2001) Crystal structure of the human nuclear cap binding complex. *Molecular cell* **8:** 383-396

McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL (1997a) 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes & development* **11:** 3306-3318

McCracken S, Fong N, Yankulov K, Ballantyne S (1997b) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*

Meinel DM, Burkert-Kautzsch C, Kieser A, O'Duibhir E, Siebert M, Mayer A, Cramer P, Söding J, Holstege FC, Sträßer K (2013) Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II. *PLoS genetics* **9**

Meinhart A (2005) A structural perspective of CTD function. *Genes & development* **19**

Merz C, Urlaub H, Will CL, Lührmann R (2007) Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *Rna*

Molenaar C, Abdulle A, Gena A, Tanke HJ (2004) Poly (A)+ RNAs roam the cell nucleus and pass through speckle domains in transcriptionally active and inactive cells. *The Journal of cell ...*

Moore MJ, Proudfoot NJ (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136:** 688-700

Mor A, Suliman S, Ben-Yishay R, Yunger S, Brody Y, Shav-Tal Y (2010) Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nature cell biology* **12:** 543-552

Mori K, Weng S-M, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Broeckhoven C (2013) The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat

proteins in FTLD/ALS. *Science* **339:** 1335-1338

Muthukrishnan S, Both GW, Furuichi Y, Shatkin AJ (1975) 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* **255:** 33-37

Neduva V, Russell RB (2006) Peptides mediating interaction networks: new leads at last. *Curr Opin Biotechnol* **17:** 465-471

Neugebauer KM (2002) On the importance of being co-transcriptional. *Journal of cell science*

Neuvéglise C, Marck C, Gaillardin C (2011) The intronome of budding yeasts. *Comptes rendus biologies* **334:** 662-670

Nojima T, Hirose T, Kimura H, Hagiwara M (2007) The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. *The Journal of biological chemistry* **282:** 15645-15651

Parker R, Sheth U (2007) P bodies and the control of mRNA translation and degradation. *Molecular cell* **25:** 635-646

Piruat JI, Aguilera A (1998) A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *The EMBO journal* **17:** 4859-4872

Politz JC, Browne ES, Wolf DE (1998) Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. *Proceedings of the ...*

Politz JC, Tuft RA, Pederson T, Singer RH (1999) Movement of nuclear poly (A) RNA throughout the interchromatin space in living cells. *Current biology*

Pryor A, Tung L, Yang Z, Kapadia F, Chang TH, Johnson LF (2004) Growth-regulated expression and G0-specific turnover of the mRNA that encodes URH49, a mammalian DExH/D box protein that is highly related to the mRNA export protein UAP56. *Nucleic acids research* **32:** 1857-1865

Reed R, Hurt E (2002) A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* **108:** 523-531

Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita V-MM, Kaivorinne A-LL, Hölttä-Vuori M, Ikonen E, Sulkava R, Benatar M, Wuu J, Chiò A, Restagno G, Borghero G, Sabatelli M, Consortium I, Heckerman D, Rogaeva E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72:** 257-268

Rodrigues JP, Rode M, Gatfield D, Blencowe BJ, Carmo-Fonseca M, Izaurralde E (2001)

REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 1030-1035

Rondón AG, Aguilera A, Struhl K, Reed R (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*

Rougemaille M, Dieppois G, Kisseleva-Romanova E, Gudipati RK, Lemoine S, Blugeon C, Boulay J, Jensen TH, Stutz F, Devaux F, Libri D (2008) THO/Sub2p functions to coordinate 3'-end processing with gene-nuclear pore association. *Cell* **135:** 308-321

Saguez C, Schmid M, Olesen J, Ghazy M, Qu X, Poulsen M, Nasser T, Moore C, Jensen T (2008) Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. *Molecular cell* **31:** 91-103

Sambrook J, Russell DW, Sambrook J (2006) *The condensed protocols from Molecular cloning : a laboratory manual*, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Schwer B, Mao X, Shuman S (1998) Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. *Nucleic acids research* **26:** 2050-2057

Schwer B, Saha N, Mao X, Chen HW, Shuman S (2000) Structure-function analysis of yeast mRNA cap methyltransferase and high-copy suppression of conditional mutants by AdoMet synthase and the ubiquitin conjugating enzyme Cdc34p. *Genetics* **155:** 1561-1576

Shabalina SA, Ogurtsov AY, Spiridonov AN, Novichkov PS, Spiridonov NA, Koonin EV (2010) Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Molecular biology and evolution* **27:** 1745-1749

Shi H, Cordin O, Minder CM, Linder P, Xu R-MM (2004) Crystal structure of the human ATP-dependent splicing and export factor UAP56. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 17628-17633

Shibuya T, Tange TO, Sonenberg N, Moore MJ (2004) eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nat Struct Mol Biol* **11:** 346-351

Shuman S (2002) What messenger RNA capping tells us about eukaryotic evolution. *Nature reviews Molecular cell biology* **3:** 619-625

Singh G, Kucukural A, Cenik C, Leszyk JD, Shaffer SA, Weng Z, Moore MJ (2012) The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151:** 750-764

Staley JP, Guthrie C (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92:** 315-326

Steckelberg A-LL, Boehm V, Gromadzka AM, Gehring NH (2012) CWC22 connects pre-mRNA splicing and exon junction complex assembly. *Cell reports* **2:** 454-461

Steckelberg AL, Gehring NH (2014) Studying the composition of mRNPs in vitro using

splicing-competent cell extracts. *Methods* **65:** 342-349

Strasser K, Hurt E (2000) Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *The EMBO journal* **19:** 410-420

Strässer K, Hurt E (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature* **413:** 648-652

Strässer K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodriguez-Navarro S, Rondón AG, Aguilera A, Struhl K, Reed R, Hurt E (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417:** 304-308

Stutz F, Bachi A, Doerks T, Braun IC, Séraphin B, Wilm M, Bork P, Izaurralde E (2000) REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA (New York, NY)* **6:** 638-650

Sugiura T, Sakurai K, Nagano Y (2007) Intracellular characterization of DDX39, a novel growth-associated RNA helicase. *Experimental cell research* **313:** 782-790

Tange TØ, Shibuya T, Jurica MS, Moore MJ (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA (New York, NY)* **11:** 1869-1883

Taniguchi I, Ohno M (2008) ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Molecular and cellular biology* **28:** 601-608

Teplova M, Wohlbold L, Khin NW, Izaurralde E, Patel DJ (2011) Structure-function studies of nucleocytoplasmic transport of retroviral genomic RNA by mRNA export factor TAP. *Nature structural & molecular biology* **18:** 990-998

Tompa P, Csermely P (2004) The role of structural disorder in the function of RNA and protein chaperones. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **18:** 1169-1175

Topisirovic I, Svitkin YV, Sonenberg N, Shatkin AJ (2011) Cap and cap-binding proteins in the control of gene expression. *Wiley interdisciplinary reviews RNA* **2:** 277-298

Tran DD, Saran S, Dittrich-Breiholz O, Williamson AJ, Klebba-Färber S, Koch A, Kracht M, Whetton AD, Tamura T (2013) Transcriptional regulation of immediate-early gene response by THOC5, a member of mRNA export complex, contributes to the M-CSF-induced macrophage differentiation. *Cell death & disease* **4**

Tutucci E, Stutz F (2011) Keeping mRNPs in check during assembly and nuclear export. *Nature reviews Molecular cell biology* **12:** 377-384

Valencia P, Dias AP, Reed R (2008) Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 3386-3391

van Dijk TB, Gillemans N, Stein C, Fanis P, Demmers J, van de Corput M, Essers J, Grosveld F, Bauer UM, Philipsen S (2010)

Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol Cell Biol* **30:** 260-272

Varia S, Potabathula D, Deng Z, Bubulya A, Bubulya PA (2013) Btf and TRAP150 have distinct roles in regulating subcellular mRNA distribution. *Nucleus* **4:** 229-240

Viphakone N, Hautbergue GM, Walsh M, Chang CT, Holland A, Folco EG, Reed R, Wilson SA (2012) TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat Commun* **3:** 1006

Visa N, Izaurralde E, Ferreira J, Daneholt B, Mattaj IW (1996) A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. *The Journal of cell biology* **133:** 5-14

Wahl MC, Will CL, Luhrmann R (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* **136:** 701-718

Wichmann I, Garcia-Lozano JR, Respaldiza N, Gonzalez-Escribano MF, Nuñez-Roldan A (1999) Autoantibodies to transcriptional regulation proteins DEK and ALY in a patient with systemic lupus erythematosus. *Human immunology* **60:** 57-62

Yamazaki T, Fujiwara N, Yukinaga H, Ebisuya M, Shiki T, Kurihara T, Kioka N, Kambe T, Nagao M, Nishida E, Masuda S (2010) The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Molecular biology of the cell* **21:** 2953-2965

Yap K, Lim ZQ, Khandelia P, Friedman B, Makeyev EV (2012) Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. *Genes & development* **26:** 1209-1223

Yap K, Makeyev EV (2013) Regulation of gene expression in mammalian nervous system through alternative pre-mRNA splicing coupled with RNA quality control mechanisms. *Molecular and cellular neurosciences* **56:** 420-428

Zenklusen D, Vinciguerra P, Wyss JC, Stutz F (2002) Stable mRNP Formation and Export Require Cotranscriptional Recruitment of the mRNA Export Factors Yra1p and Sub2p by Hpr1p. *Molecular and Cellular Biology*

Zhang M, Green MR (2001) Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56. *Genes & development*

Zhang Z, Krainer AR (2007) Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components. *Proceedings of the National Academy of Sciences* **104**

Zhou Z, Luo M, Straesser K, Katahira J (2000) The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature*

Acknowledgments

First of all I would like to thank my supervisor – Niels Gehring without whom it would not be possible to start and finish this thesis. He supported me throughout this long, scientific journey with discussions, inspiration and advice.

Moreover, I would like to thank Prof. Dr. Karin Schnetz for her time to read and grade this thesis, Prof. Dr. Ulrich Baumann for being part of my thesis committee.

I would also like to mention a great amount of help from the Niels Gehring Lab Members. My colleagues that became friends – Ben, Franzi, Kusum, Lena, Tobi, Simona – thank you for all these years of support. We are a great team. Special thanks to Volker, it was a pleasure to share a lab space for these four years with you.

I would like to thank Heidi and Juliane for patience and excellent technical support.

I would like to mention members of Mirka Uhlirova Lab for sharing their lab equipment and expertise.

This thesis and my stay in Cologne was supported by the IGS DHD Team. Special thanks for Dr. Isabell Witt for creating a friendly environment for me here, at the Institute for Genetics.

Many thanks to my best friends that hold my back constantly – Antonella, Eva and Simona. Special thanks to Ania and Michał for our discussions and support.

It is my honour to dedicate this work to my family – Tą pracę dedykuję Rodzicom i Siostrze.

Hvala za vse, Jernej.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit–einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von PD Dr. Niels H. Gehring betreut worden.

Köln, den 24.11.2014

Agnieszka Gromadzka

Curriculum Vitae

Agnieszka Marta Gromadzka

Born: 10th April 1986 in Poznań, Poland

Current address: Paul Schallück Str. 29, 50939 Cologne, Germany

Higher education

2010-2014 **PhD student**

Thesis title: Identification of determinants for export factor binding to spliced RNAs

Supervisor: PhD. N.H. Gehring, Institute of Genetics, University of Cologne, Germany

2008-2010 **M.Sc. in Biotechnology**; with honours

Thesis title: Human chorionic gonadotropin beta subunit expression causes increased proliferation of cancer cells
Thesis supervisor: PhD. Anna Jankowska, University of Medical Sciences, Poznań, Poland

2005-2008 **B.Sc. in Biotechnology**;
Supervisor: Prof. PhD. Anna Goździcka-Józefiak, Adam Mickiewicz University, Poznań

Highschool education

2001-2005 I Liceum Ogólnoszktałcące, Swarzędz, Poland